INTERSCIENCE TRACTS
IN PURE AND APPLIED MATHEMATICS
NUMBER 11

SERGE LANG

•

DIOPHANTINE GEOMETRY

INDIANA UNIVERSITY
NORTHWEST

LIBRARY

Digitized by the Internet Archive
in 2023 with funding from
Kahle/Austin Foundation

https://archive.org/details/diophantinegeome0000serg

INTERSCIENCE TRACTS IN PURE AND APPLIED MATHEMATICS

Editors: L. BERS · R. COURANT · J. J. STOKER

1. *D. Montgomery and L. Zippin* - **Topological Transformation Groups**
2. *Fritz John* - **Plane Waves and Spherical Means Applied to Partial Differential Equations**
3. *E. Artin* - **Geometric Algebra**
4. *R. D. Richtmyer* - **Difference Methods for Initial-Value Problems**
5. *Serge Lang* - **Introduction to Algebraic Geometry**
6. *Herbert Busemann* - **Convex Surfaces**
7. *Serge Lang* - **Abelian Varieties**
8. *S. M. Ulam* - **A Collection of Mathematical Problems**
9. *I. M. Gel'fand* - **Lectures on Linear Algebra**
10. *Nathan Jacobson* - **Lie Algebras**
11. *Serge Lang* - **Diophantine Geometry**

Additional volumes in preparation

INTERSCIENCE TRACTS
IN PURE AND APPLIED MATHEMATICS

Editors: L. BERS · R. COURANT · J. J. STOKER

Number 11

DIOPHANTINE GEOMETRY

By Serge Lang

INTERSCIENCE PUBLISHERS
a division of John Wiley & Sons, New York · London

DIOPHANTINE GEOMETRY

SERGE LANG
Columbia University, New York, New York

INTERSCIENCE PUBLISHERS
a division of John Wiley & Sons, New York · London

INDIANA
UNIVERSITY
LIBRARY

NORTHWEST

COPYRIGHT © 1962 BY JOHN WILEY & SONS, INC.

ALL RIGHTS RESERVED

LIBRARY OF CONGRESS CATALOG CARD NUMBER 62-16309

Printed in the United States of America

Foreword

Diophantine problems represent some of the strongest aesthetic attractions to algebraic geometry. They consist in giving criteria for the existence of solutions of algebraic equations in rings and fields, and eventually for the number of such solutions.

The fundamental ring of interest is the ring of ordinary integers **Z**, and the fundamental field of interest is the field **Q** of rational numbers. One discovers rapidly that to have all the technical freedom needed in handling general problems, one must consider rings and fields of finite type over the integers and rationals. Furthermore, one is led to consider also finite fields, $\mathfrak{p}$-adic fields (including the real and complex numbers) as representing a localization of the problems under consideration.

We shall deal with global problems, all of which will be of a qualitative nature. On the one hand we have curves defined over say the rational numbers. If the curve is affine one may ask for its points in **Z**, and thanks to Siegel, one can classify all curves which have infinitely many integral points. This problem is treated in Chapter VII. One may ask also for those which have infinitely many rational points, and for this, there is only Mordell's conjecture that if the genus is $\geqq 2$, then there is only a finite number of rational points.

On the other hand, we have abelian varieties. If A is an abelian variety defined over **Q**, then its group of rational points is finitely generated (Mordell-Weil theorem). The proofs do not give a constructive method to get hold of the generators. Abelian varieties include of course curves of genus 1, which have been the fundamental testing ground for conjectures, theorems, and proofs in the theory.

As a curve of genus $\geqq 1$ is characterized by the fact that it is embeddable in an abelian variety, one is led to study subvarieties of abelian varieties, and especially Mordell's conjecture, in this light. A curve of genus $\geqq 2$ is characterized by the fact that it is unequal to any non-zero translation of itself in its Jacobian. This is a very difficult hypothesis to use.

It is also interesting to look at subvarieties of other group varieties. For us, the natural ones to consider are the toruses, i.e. products of multiplicative groups. It will be shown that if G is such a group, V a curve contained in G, Γ a finitely generated subgroup of G, and $V \cap \Gamma$ is infinite then V is the translation of a subtorus (in characteristic 0).

For various generalizations, and comments on the above results and problems, see Chapters VI and VII and the Historical Notes at the end of these chapters.

The prototypes of the diophantine results and methods given in this book were developed by Mordell, Weil, and Siegel between 1920 and 1930. Considering recent developments in the techniques of algebraic geometry, and Roth's theorem, it seemed quite worth while to give a systematic exposition of the theory as it stands, i.e. essentially a qualitative theory. One of the directions of research at present lies in making theorems quantitative, that is, getting estimates for the number of rational or integral points, or the generators of the Mordell-Weil group. It is a striking feature of the proofs that they are highly non-constructive. There do exist still some unsolved qualitative problems. Aside from Mordell's conjecture, for instance, I would conjecture that Siegel's result on integral points remains true for an affine open subset of an abelian variety, defined over **Z** (or a ring of finite type over **Z**). In general, the extension of Roth's theorem and the theory of integral points as given in Chapter VII to varieties of dimension >1 is still lacking.

New York, 1961 S. Lang

Prerequisites

They are of different type, some of the chapters being very elementary, and others requiring more background.

Chapters I, II and Chapter III, §1, §2 are elementary and self contained, assuming only a knowledge of elementary algebra, say at the level of Galois theory, together with a few facts on valuation rings which may be found in Chapter I of [16]. Chapter III, §3 requires some knowledge of algebraic geometry, of an essentially standard nature.

Chapter IV depends only on elementary facts concerning absolute values, together with the notion of linear systems and their basic properties as exposed in [16], Chapter VI, except for §3 which demands a more extensive knowledge, including some facts on abelian varieties. However, in our applications, we need the result of Chapter IV, §3 only for curves (in which case it can be proved by using the Riemann-Roch theorem) and for abelian varieties in which case it is easily reduced to the criterion of §2.

Chapter V uses the theory of reduction mod $\mathfrak{p}$ and the theory of abelian varieties. The reader will find all necessary background in [36] or [11].

Chapter VI is again elementary, depending only on simple properties of absolute values.

Chapter VII is axiomatized to the greatest possible extent, but its results combine everything that precedes it.

Chapter VIII is again almost entirely elementary and in particular, the proof of Hilbert's theorem for the rationals can be read with only a knowledge of beginning calculus.

The reader will therefore see that Chapters I and II give a good part of classical algebraic number theory. Chapters III (except for §3) and VI may be viewed as providing a continuation of this algebraic number theory in a more specialized direction, distinct from the multiplicative direction which may be taken by going into class field theory at that point.

For an introduction to algebraic geometry, I refer the reader to [16], which will be quoted as IAG. For abelian varieties, all results needed are contained in [17], which will be referred to as

AV, with the exception of properties of reduction mod $\mathfrak{p}$ (cf. [36] or [11]). References to Weil's *Foundations* will be to the first or second edition, in canonical style $F^1 - X_y$ Th. Z, or $F^2 - X_y$ Th. Z.

A list of symbols used most frequently will be found on p. 167. Unless otherwise specified, all rings are without zero divisors, and with unit elements. **Z** is the integers, and **Q** the rational numbers.

CONTENTS

CHAPTER I

Absolute Values

1. Definitions, dependence and independence 1
2. Completions 5
3. Unramified extensions 9
4. Finite extensions 13

CHAPTER II

Product Formula

1. Proper sets of absolute values 19
2. Divisors, ideals, and units 25
3. Unit theorem 30
4. Relative units and ideal classes 37

CHAPTER III

Heights

1. Definitions 43
2. Gauss' lemma 45
3. Heights in function fields 50
4. Heights on abelian groups 54

CHAPTER IV

Properties of Heights

1. Functorial properties 61
2. Criterion for equivalence 64
3. Criterion for quasi-equivalence 67
4. Projections on curves 69

CHAPTER V

The Mordell-Weil Theorem

1. Reduction steps 72
2. Kummer theory 75
3. The weak Mordell-Weil theorem 80
4. The infinite descent 81
5. Points of bounded height 82
6. Theorem of the base 86

CHAPTER VI

The Thue-Siegel-Roth Theorem

1. Statement of the theorem 91
2. Reduction to simultaneous approximations 96
3. Sketch of the proof 98
4. A combinatorial lemma 102
5. Proof of Proposition 1 104
6. Wronskians 106
7. Factorization of a polynomial 108
8. The index 110
9. Proof of Proposition 2 113
10. A geometric formulation of Roth's theorem 115

CHAPTER VII

Siegel's Theorem

1. Height of integral points 122
2. Finiteness theorems 125
3. Theorem of decomposition 128
4. Curves of genus 0 134

CHAPTER VIII

Hilbert's Irreducibility Theorem

1. Irreducibility and integral points 142
2. Irreducibility over the rational numbers 146
3. Reduction steps 150
4. Function fields 152
5. Abstract definition of Hilbert sets 155
6. Applications to commutative group varieties 158

Bibliography 163

Table of Notation 167

Index 169

CHAPTER I

Absolute Values

What distinguishes an arithmetic problem from a geometric one is the nature of the ground ring or ground field and any prime structure it may have. In addition to the existence of primes (prime numbers, prime ideals, etc.) one must also, or even principally, take into account archimedean absolute values. It turns out that all the foundational material we need can be stated in terms of absolute values, and we have included in this chapter everything we need in the sequel.

It is convenient to localize certain questions by taking the completion of our field with respect to an absolute value. In fact, our insight concerning the extensions of an absolute value on K to a finite extension E will comes form an embedding of E into the algebraic closure of the completion of K.

§ 1. *Definitions, dependence and independence*

Let K be a field. An *absolute value* on K is a real valued function $x \to v(x)$ on K satisfying the following three properties:

AV 1. We have $v(x) \geqq 0$, and $v(x) = 0$ if and only if $x = 0$.

AV 2. $v(xy) = v(x)v(y)$ for all $x, y \in K$.

AV 3. $v(x + y) \leqq v(x) + v(y)$.

If instead of AV 3, the absolute value satisfies the stronger condition

AV 4. $v(x + y) \leqq \max [v(x), v(y)]$

then we shall say that it is a *valuation*, or that it is *non-archimedean*. In this chapter we emphasize absolute values, taking more or less for granted the elementary properties of valuations as given, for instance, in IAG [16], Chapter I.

The absolute value which is such that $v(x) = 1$ for all $x \neq 0$ is called *trivial*.

An absolute value on K defines a metric, and thus a topology. We shall therefore write $|x|_v$ or simply $|x|$ instead of $v(x)$. Two absolute values are called *dependent* if they define the same topology.

If they do not, they are called *independent.*

PROPOSITION 1. *Two non-trivial absolute values, $| \ |_1$ and $| \ |_2$, on a field K, are dependent if and only if the relation*

$$| x |_1 < 1$$

implies $| x |_2 < 1$. If they are dependent, then there exists a real number $\lambda > 0$ such that $| x |_1 = | x |_2^{\lambda}$ for all $x \in K$.

Proof: If the two absolute values are dependent, then our condition is satisfied, because the set of $x \in K$ such that $| x |_1 < 1$ is the same as the set such that $\lim x^n = 0$ for $n \to \infty$. Conversely, assume the condition satisfied. Then $| x |_1 > 1$ implies $| x |_2 > 1$ since $| x^{-1} |_1 < 1$. By hypothesis, there exists an element $x_0 \in K$ such that $| x_0 |_1 > 1$. Put $a = | x_0 |_1$ and $b = | x_0 |_2$, and let $\lambda = \log(b)/\log(a)$. Let $x \in K$, and $x \neq 0$. Then $| x |_1 = | x_0 |_1^{\alpha}$ for some $\alpha > 0$. If m, n are integers > 0 such that $m/n > a$ we have $| x |_1 < | x_0 |_1^{m/n}$ whence $| x^n/x_0^m |_1 < 1$, and thus $| x^n/x_0^m |_2 < 1$, or in other words, $| x |_2 < | x_0 |_2^{m/n}$. Similarly, $| x |_2 > | x_0 |_2^{m/n}$. Hence $| x |_2 = | x_0 |_2^{\alpha}$. The assertion of our proposition is now obvious.

One of the main applications of the theory of absolute values will be to the rational numbers **Q** and its finite extensions (which are called *number fields*). It is easy to show that a non-trivial absolute value on **Q** is dependent on one of the following:

The ordinary absolute value.

The *p-adic absolute value*, $v_p = | \ |_p$, defined for each prime number p by the formula

$$| p^r m/n |_p = 1/p^r$$

where r is an integer, and m, n are integers $\neq 0$ and not divisible by p.

In fact, if the absolute value on **Q** is a valuation, then looking at its valuation ring $\mathfrak{o}$ and maximal ideal $\mathfrak{m}$ one sees immediately that $\mathfrak{m} \cap \mathbf{Z}$ is a prime ideal, and thus is generated by a prime number p. Thus by Proposition 1 our absolute value is a power of the p-adic absolute value (cf. IAG Chapter I).

Suppose that we have an absolute value on a field which is bounded on the prime ring (*i.e.* the integers **Z** in characteristic 0, or the integers mod p in characteristic p). Then it is non-archimedean. One sees this immediately by looking at the expansion of $(x + y)^n$ and $| (x + y)^n) |^{1/n}$ as $n \to \infty$. This is always the case in characteristic > 0.

If the absolute value is archimedean, then we refer the reader to [3] or [39] for a proof that it is dependent on the ordinary absolute value. We do not need this fact in our applications, for we always will start with a concretely given set of absolute values on fields which will interest us.

The relation of dependence on absolute values is obviously an equivalence relation (symmetric, reflexive, transitive). An equivalence class of non-trivial absolute values on a field K will be called a *prime of* K, and will sometimes be denoted by a German letter $\mathfrak{p}$, or if we deal with an extension of K, by $\mathfrak{P}$. If E is an extension of K, and $\mathfrak{p}$ a prime of K, and $\mathfrak{P}$ a prime of E, we say that $\mathfrak{P}$ *extends* $\mathfrak{p}$ and write $\mathfrak{P} \mid \mathfrak{p}$ if the absolute values in $\mathfrak{P}$ when restricted to K give absolute values of $\mathfrak{p}$. Similarly, if w is an absolute value on E extending an absolute value on K, we write $w \mid v$. We shall study in § 4 the manner in which an absolute value on K can be extended to algebraic extensions of K.

For the convenience of the reader, we recall some terminology. A prime $\mathfrak{p}$ is called *non-archimedean* if its absolute values are non-archimedean. If this is the case, then it belongs to a valuation ring $\mathfrak{o} = \mathfrak{o}_{\mathfrak{p}}$ with maximal ideal $\mathfrak{m} = \mathfrak{m}_{\mathfrak{p}}$. Let E be a finite extension of K, and let $\mathfrak{P}$ be a prime extending $\mathfrak{p}$. It corresponds to a valuation ring lying above $\mathfrak{o}$. Let $\mathfrak{o}_E$ be this valuation ring, and $\mathfrak{m}_E$ its maximal ideal. Then the residue class field $\mathfrak{o}_E/\mathfrak{m}_E$ is an extension of $\mathfrak{o}/\mathfrak{m}$, and its degree is called the *residue class degree*, denoted by $d_{\mathfrak{P}}$, or d_w if w is one of the valuations of $\mathfrak{P}$. The value group $w(K^*)$ is a subgroup of $w(E^*)$, whose index is called the *ramification index* and denoted by $e_{\mathfrak{P}}$ or e_w. They are both finite (cf. IAG Chapter I).

In our applications we shall deal mostly with archimedean absolute values, or with discrete valuations. Let us say a few words about the latter.

Suppose we have a field K with a valuation ring $\mathfrak{o}$ and maximal ideal $\mathfrak{p}$. In general, we obtain a valuation whose canonical value group is the factor group K^*/U where U are the units of $\mathfrak{o}$ (cf. IAG, Chapter I). Suppose this value group is infinite cyclic. It can then be embedded into the positive reals, preserving the ordering. We select a canonical generator for it (from two possible ones) namely, $|\pi|$ where π lies in $\mathfrak{p}$. Then every element of K^* can be written in the form

$$x = u\pi^r$$

where u is a unit, r an integer, which is called the *order* of x at the valuation (or prime). In particular, $\mathfrak{p}$ is principal, generated by π. Using the same letter $\mathfrak{p}$ for the maximal ideal of $\mathfrak{o}$, or for the equivalence class of absolute values (valuations) which correspond to it, we write

$$r = \operatorname{ord} x = \operatorname{ord}_{\mathfrak{p}} x = \operatorname{ord}_v x .$$

Using Proposition 1, we see that all the absolute values corresponding to $\mathfrak{o}$ can be described as follows. We choose a number c c such that $0 < c < 1$, and put

$$| x | = c^{\operatorname{ord} x} .$$

Such absolute values, or valuations, will be called *discrete.*

There is of course considerable arbitrariness in the choise of the constant c, and there is usually no way of choosing it in a natural manner. However, in some applications, the residue class field $\mathfrak{o}/\mathfrak{p}$ is a finite field. In that case, denoting the number of elements in it by $\mathbf{N}\mathfrak{p}$, we choose $c = 1/\mathbf{N}\mathfrak{p}$, and our absolute value becomes

$$\| x \|_{\mathfrak{p}} = (1/\mathbf{N}\mathfrak{p})^{\operatorname{ord}_{\mathfrak{p}} x} .$$

This will be the case in one of the most important applications, namely, number fields, whose non-trivial valuation rings all have a finite residue class field. We shall return to this in greater detail in Chapter II.

In Proposition 1 we derived a strong condition on dependent absolute values. We shall now derive a condition on independent ones.

THEOREM 1. *Let K be a field, and $| \ |_1, \cdots, | \ |_s$ non-trivial pairwise independent absolute values on K. Let $x_1, \cdots, x_s$ be elements of K, and $\varepsilon > 0$. Then there exists $x \in K$ such that*

$$| x - x_i |_i < \varepsilon$$

for all i.

Proof: Consider first two of our absolute values, say v_1 and v_s. By hypothesis, we can find $\alpha \in K$ such that $| \alpha |_1 < 1$ and $| \alpha |_s \geqq 1$. Similarly, we can find $\beta \in K$ such that $| \beta |_1 \geqq 1$ and $| \beta |_s < 1$. Put $y = \beta/\alpha$. Then $| y |_1 > 1$ and $| y |_s < 1$.

We shall now prove that there exists $z \in K$ such that $|z|_1 > 1$ and $|z|_j < 1$ for $j = 2, \cdots, s$. We prove this by induction, the case $s=2$ having just been proved. Suppose we have found $z \in K$ satisfying

$$|z|_1 > 1 \qquad \text{and} \qquad |z|_j < 1 \qquad \text{for} \quad j = 2, \cdots, s-1 .$$

If $|z|_s \leqq 1$, then the element $z^n y$ for large n will satisfy our requirements.

If $|z|_s > 1$, then the sequence

$$t_n = z^n/(1 + z^n)$$

tends to 1 at v_1 and v_s, and tends to 0 at v_j $(j = 2, \cdots, s-1)$. For large n, it is then clear that $t_n y$ satisfies our requirements.

Using the element z that we have just constructed, we see that the sequence $z^n/(1 + z^n)$ tends to 1 at v_1 and to 0 at v_j for $j = 2, \cdots, s$. For each $i = 1, \cdots, s$ we can therefore construct an element z_i which is very close to 1 at v_i and very close to 0 at v_j $(j \neq i)$. The element

$$x = z_1 x_1 + \cdots + z_s x_s$$

then satisfies the requirement of the theorem.

§ 2. *Completions*

Let K be a field with a non-trivial absolute value v, which will remain fixed throughout this section. One can then define in the usual manner the notion of Cauchy sequence, and K is said to be complete if every Cauchy sequence converges.

Proposition 2. *There exists a pair (K_v, i) consisting of a field K_v complete under an absolute value, and an embedding $i\colon K \to K_v$ such that the absolute value on K is induced by that of K_v* (i. e. $|x|_v = |ix|$), *and such that iK is dense in K_v. If (K'_v, i') is another such pair, then there exists a unique isomorphism $\varphi\colon K_v \to K'_v$ preserving the absolute values, and making the following diagram commutative*:

$$\begin{tikzcd}
K_v \arrow[rr, "\varphi"] & & K'_v \\
& K \arrow[ul, "i"] \arrow[ur, "i'"'] &
\end{tikzcd}$$

The uniqueness is obvious. One proves the existence in the well known manner: One considers the Cauchy sequences in K, which form a ring. The null sequences form a maximal ideal, and the residue class ring is a field K_v. Our field K is naturally embedded in K_v (by means of the sequences whose elements consist of a fixed element of K), and the absolute value on K can be extended to K_v by continuity.

One usually identifies K inside K_v, and calls K_v the *completion* of K.

If the absolute value is archimedean (and thus K has characteristic 0), then it is well known that K_v is either the field of real numbers or of complex numbers, and that the absolue value is dependent on the ordinary one. For a proof, the reader may see [3]. We shall not need this result in this book, and hence we shall always *assume* that the archimedean absolute values we deal with have this property.

As the completion depends only on the prime determined by our absolute value, we also denote it by $K_{\mathfrak{p}}$.

If the absolute value is a valuation, then we refer the reader to IAG, Chapter I, for the proof that it can be extended in some way to the algebraic closure of K.

Proposition 3. *Let K be complete with respect to a non-trivial absolute value v. If E is any algebraic extension of K, then v has a unique extension to E. If E is finite over K, then E is complete.*

Proof: We give the proof only in the non-archimedean case, the other being trivial once one knows that the only complete archimedean fields are the reals or complex. Without loss of generality, we may assume E finite over K. Let us prove the uniqueness. In view of Proposition 1, it suffices to prove that two extensions define the same topology. Furthermore, if $\omega_1, \cdots, \omega_n$ are linearly in dependent elements of E over K, and

$$x^{(\nu)} = \xi_1^{(\nu)}\omega_1 + \cdots + \xi_n^{(\nu)}\omega_n \qquad \xi_n^{(\nu)} \in K \tag{1}$$

is a Cauchy sequence in E, then it will suffice to prove that the n sequences $\xi_j^{(\nu)}$ converge in K. We do this by induction on n. It is obvious for $n = 1$. Assume $n \geqq 2$. We consider a sequence as above and, without loss of generality, we may assume that it converges to 0. (If necessary, consider $x^{(\nu)} - x^{(\mu)}$ for $\nu, \mu \to \infty$.) We must then show that the sequences of the coefficients converge to 0 also. If this is not the case, then there exists a number $a > 0$ such that we have for some j, say $j = 1$,

$$|\xi_1^{(\nu)}| > a$$

for all ν sufficiently large. Then $x^{(\nu)}/\xi_1^{(\nu)}$ converges to 0, and we can write

$$\frac{x^{(\nu)}}{\xi_1^{(\nu)}} - \omega_1 = \frac{\xi_2^{(\nu)}}{\xi_1^{(\nu)}}\omega_2 + \cdots + \frac{\xi_n^{(\nu)}}{\xi_1^{(\nu)}}\omega_n .$$

We let $y^{(\nu)}$ be the right-hand side of this equation. Then $y^{(\nu)}$ yields a sequence which converges (according to the left-hand side of the equation). By induction, we conclude that its coefficients in terms of $\omega_2, \cdots, \omega_n$ also converge in K, say to $\eta_2, \cdots, \eta_n$. Taking the limit, we get

$$\omega_1 = \eta_2\omega_2 + \cdots + \eta_n\omega_n$$

contradicting the independence of the ω_j.

We have thus proved that the extension of our valuation is unique. Furthermore, we have seen that if E is finite over K, of degree n, then it has the topology of Cartesian n-space over K: in particular, it is complete. The Cauchy sequences of the coefficients in (1) above converge in K, say to $\xi_1, \cdots, \xi_n$; then our sequence $x^{(\nu)}$ converges to $\xi_1\omega_1 + \cdots + \xi_n\omega_n$.

From the uniqueness, we can get an explicit determination of the absolute value on an algebraic extension. Indeed, observe first that if E is a normal extension of K, and σ is an automorphism of E over K, then the function

$$x \to |\sigma x|$$

is an absolute value on E extending that of K. Hence we get

$$|\sigma x| = |x| .$$

If E is algebraic over K, and σ is an isomorphism of E over K, then the same conclusion remains valid, as one sees immediately by embedding E in a normal extension of K. (We assume, without loss of generality, that all our algebraic extensions are contained in a fixed algebraic closure of K.) In particular, if α is algebraic over K, of degree n, and if $\alpha_1, \cdots, \alpha_n$ are its conjugates (counting multiplicities), then all the absolute values $|\alpha_i|$ are equal. Denoting by N the norm from $K(\alpha)$ to K, we see that

$$|N(\alpha)| = |\alpha|^n ,$$

and taking the n-th root, we get:

Proposition 4. *Let K be complete with respect to a non-trivial* *absolute value. Let* α *be algebraic over* K, *and let* N *be the norm from* $K(\alpha)$ *to* K. *Let* $n = [K(\alpha) : K]$. *Then*

$$|\alpha| = |N(\alpha)|^{1/n} .$$

Of course, the arguments we gave above hold when we deal with a valuation, but in the archimedean case the proposition is also true, taking into account that we are dealing with the complex and real numbers.

Let us now consider the case of a discrete valuation. We assume that K is complete under a discrete valuation, with valuation ring $\mathfrak{o}$, and maximal ideal $\mathfrak{m}$, and we let π be a prime element of $\mathfrak{o}$. If $x, y \in K^*$ are such that $|x| = |y|$, we shall write $x \sim y$. Let π_i ($i = 1, 2, \cdots$) be a sequence of elements of $\mathfrak{o}$ such that $\pi_i \sim \pi^i$. If $x \in \mathfrak{o}$ we denote by x' its canonical image in $\mathfrak{o}/\mathfrak{m}$. Let R be a set of representatives of $\mathfrak{o}/\mathfrak{m}$ in $\mathfrak{o}$. This means that the map $x \to x'$ of R into $\mathfrak{o}/\mathfrak{m}$ is bijective. Then every element of $\mathfrak{o}$ can be written as a convergent series

$$x = a_0 + a_1\pi_1 + a_2\pi_2 + \cdots$$

with $a_i \in R$, and the a_i uniquely determined. This is easily seen by a recursive argument, namely, suppose we have written

$$x \equiv a_0 + \cdots + a_n\pi_n \pmod{\mathfrak{m}^{n+1}}$$

then $x - (a_0 + \cdots a_n\pi_n) = \pi_{n+1}y$ with some $y \in \mathfrak{o}$. But by hypothesis, we can write $y = a_{n+1} + \pi z$ with some $a_{n+1} \in R$. From this we get

$$x \equiv a_0 + \cdots + a_{n+1}\pi_{n+1} \pmod{\mathfrak{m}^{n+2}} ,$$

and it is clear that the n-th term in our series tends to 0. Thus our series converges. The fact that R contains precisely one representative for each residue class mod $\mathfrak{m}$ implies that the a_i are uniquely determined.

Let E be a finite extension of K, let $\mathfrak{o}_E$, $\mathfrak{m}_E$ be the valuation ring and maximal ideal in E lying above $\mathfrak{o}$, $\mathfrak{m}$ in K, and let Π be a prime element in E. If Γ_E and Γ_K are the value groups of the valuation in E and K respectively, let $e = (\Gamma_E : \Gamma_K)$. It is the *ramification index*. We then have

$$|\Pi^e| = |\pi| ,$$

and the elements

$$\Pi^i\pi^j \qquad \begin{array}{l} 0 \leqq i \leqq e - 1 \\ j = 0, 1, 2, \cdots \end{array}$$

have order $je+i$ in E.

Let $\omega_1, \cdots, \omega_d$ be elements of E such that their residue classes $\omega_1', \cdots, \omega_d'$ in $E' = \mathfrak{o}_E/\mathfrak{m}_E$ form a basis for E' over K'. If R is as above a set of representatives of $\mathfrak{o}/\mathfrak{m}$ in $\mathfrak{o}$, then

$$a_1\omega_1 + \cdots + a_d\omega_d\ ,$$

as the a_ν range over R, form a system of representatives of $\mathfrak{o}_E/\mathfrak{m}_E$ by definition. From this, one sees that every element of $\mathfrak{o}_E$ admits a convergent expansion

$$\sum_{i=0}^{e-1} \sum_{\nu=1}^{d} \sum_{j=0}^{\infty} a_{\nu,i,j}\pi^j \omega_\nu \Pi^i$$

which shows that the ed elements $\omega_\nu \Pi^i$ form a set of generators of $\mathfrak{o}_E$ over $\mathfrak{o}$. On the other hand, one sees easily that they are linearly independent over K, and hence we get:

Proposition 5. *Let K be complete with respect to a discrete valuation. Let E be a finite extension of K, with ramification index e and residue class degree d. Then*

$$ed = [E:K]\ .$$

If we now combine the interpretation of the degree given in Proposition 5 for discrete valuations with the expression given in Proposition 4 we get:

Proposition 6. *Let K be complete with respect to a discrete valuation and let $\mathfrak{p}$ be its prime. Let E be a finite extension of K, $\mathfrak{P}$ the prime extending $\mathfrak{p}$, and $d_\mathfrak{P}$ the degree of the residue class field extension. Then for $\alpha \in E^*$, we have*

$$\operatorname{ord}_\mathfrak{p} N_K^E(\alpha) = d_\mathfrak{P} \operatorname{ord}_\mathfrak{P} \alpha\ .$$

Proof: This is immediate from the formula

$$|\, N_K^E(\alpha)\,| = |\,\alpha\,|^{e_\mathfrak{P} d_\mathfrak{P}}$$

and the definitions.

§ 3. *Unramified extensions*

Let K be a field, $\mathfrak{o}$ a valuation ring of K, and $\mathfrak{m}$ its maximal ideal. We let $K' = \mathfrak{o}/\mathfrak{m}$ be the residue class field, and we let φ_0 be the canonical K'-valued placed on K. We assume that φ_0 is extended in a fixed way to the algebraic closure $\bar{K}$ of K. and call φ this extension.

Let E be an algebraic extension of K, and $\mathfrak{o}_E$ one of the valuation rings of E lying above $\mathfrak{o}$, with $\mathfrak{m}_E$ as maximal ideal. If E is finite over K, then $\mathfrak{o}_E/\mathfrak{m}_E$ is finite over $\mathfrak{o}/\mathfrak{m}$, and thus φ is $\bar{K}'$-valued. We frequently identify $\varphi(\mathfrak{o}_E)$ with $\mathfrak{o}_E/\mathfrak{m}_E$, and denote it by E'.

Let v be a valuation associated with $\mathfrak{o}$, and w an extension of v to E associated with $\mathfrak{o}_E$. We shall say that w is *unramified* over v, or simply unramified, if the residue class field extension E' over K' is separable, and if the ramification index is equal to 1, i.e. if

$$(w(E^*): v(K^*)) = 1 \ .$$

If every w extending v is unramifed we say that v is *unramified.*

The purpose of this section is to study unramified extensions when K is complete. However, we may just as well make a weaker assumption:

Throughout the rest of this section, we shall assume that for each algebraic extension E of K, there exists precisely one valuation ring of E lying above $\mathfrak{o}$.

In view of our assumption, we know that if ψ is another place of $\bar{K}$ having the same restriction as φ on K, then ψ and φ are equivalent, i.e. there exists an automorphism σ' of $\bar{K}'$ over K' such that

$$\psi = \sigma'\varphi \ .$$

If $f(X)$ is a polynomial in $\bar{K}[X]$ whose coefficients are finite under φ, then we let f^φ be the polynomial obtained by applying φ to its coefficients. Suppose f has leading coefficient 1 and that its coefficients are integral over $\mathfrak{o}$. If

$$f(X) = \prod_{i=1}^{d} (X - \alpha_i)$$

is a factorization in $\bar{K}$, then

$$f^\varphi(X) = \prod_{i=1}^{d} (X - \varphi\alpha_i)$$

is a factorization of f^φ in the algebraic closure of K'.

If $f'(X)$ is a polynomial in $E'[X]$ then any polynomial f in $\mathfrak{o}_E[X]$ of the same degree as f', with leading coefficient 1 if f' has leading coefficient 1, and such that $f^\varphi = f'$ will be said to be *lifted back from* f' *into* E. It is clearly always possible to lift back a polynomial.

The above remark concerning the factorization of f shows that φ maps $\bar{K}$ onto the algebraic closure of K'.

Our assumption concerning the uniqueness of the extension of the valuation ring yields the following result.

Proposition 7. *Let f be a polynomial in $\mathfrak{o}[X]$ such that f is irreducible in $K[X]$ and has leading coefficient* 1. *Then f^φ is the power of an irreducible polynomial in $K'[X]$.*

Proof: Suppose $f^\varphi = g'h'$ where g', h' are relatively prime of degrees $\geqq 1$. Suppose we have a factorization as above, and say $\varphi\alpha_1$ is a root of g' and $\varphi\alpha_2$ is a root of h'. There exists an isomorphism σ of $K(\alpha_1)$ over K such that $\sigma\alpha_1 = \alpha_2$ because f is irreducible, and we can extend σ to $\bar{K}$. Thus $\varphi\sigma\alpha_1 = \varphi\alpha_2$. But $\varphi\sigma$ is a place coinciding with φ on K, and thus must be conjugate to φ, i.e. there exists an isomorphism σ' of $\bar{K}'$ over K' such that $\varphi\sigma = \sigma'\varphi$, and we must have $\sigma'\varphi\alpha_1 = \varphi\alpha_2$. This contradicts the fact that g' and h' are relatively prime, and proves that f^φ is the power of an irreducible polynomial.

Our purpose is now to establish a bijective correspondence between certain extensions of K and separable extensions of K'. We shall say that a finite extension E of K is *unramified* if it satisfies the following properties:

UN 1. The residue class field E' is separable over K'.

UN 2. We have $[E:K] = [E':K']$.

Another way of formulating these conditions is:

Proposition 8. *Let E be a finite extension of K, and let φ_0 be the canonical place of K. Let $[E:K] = n$. Then E is unramified over K if and only if φ_0 has at least n distinct extensions to places of E (in the given algebraic closure of K'), and in that case, it has exactly n.*

Proof: By uniqueness, all extensions of φ_0 to E are conjugate, and the number of conjugates is equal to the separable degree of E' over K'. As $[E':K'] \leqq [E:K]$ our assertion is immediate.

We observe that if E is unramified over K, then the ramification index is equal to 1. In the case of greatest interest to us, this is characteristic.

Proposition 9. *Assume K complete under a discrete valuation. Then the finite extension E of K is unramified if and only if E' is separable over K' and the ramification index is equal to* 1.

Proof: This follows immediately from Proposition 5.

An arbitrary algebraic extension of K is unramified if every finite subextension is unramified.

We shall now obtain an algebraic criterion for non-ramification.

Proposition 10. *Let E be finite over K. If E is unramified, let $\alpha' \in E'$ be such that $E' = K'(\alpha')$, and let $\alpha \in E$ be such that $\varphi\alpha = \alpha'$. Then $E = K(\alpha)$, and the irreducible polynomial $f(X)$ of α over K is such that f^φ is irreducible. Conversely, if $K = K(\alpha)$ for some integer α satisfying a polynomial $f(X)$ in $\mathfrak{o}[X]$ having leading coefficient 1 and such that f^φ has no multiple root, then E is unramified over K and $E' = K'(\varphi\alpha)$.*

Proof: First assume E unramified. Let $f'(X)$ be the irreducible polynomial of α' over K'. Let α be an element of E such that $\varphi\alpha = \alpha'$, and let $f(X)$ be its irreducible polynomial over K. Then α is integral over $\mathfrak{o}$, and α' is a root of f^φ, and hence f' divides f^φ. On the other hand,

$$\deg f' = [E' : K'] = [E : K] \geqq \deg f$$

and so $f' = f^\varphi$ This proves the first statement.

Conversely, if α satisfies the stated condition, then we may assume without loss of generality that its irreducible polynomial $g(X)$ is such that g^φ has no multiple roots, because g divides f. We can now simply apply Proposition 7, to conclude that g^φ is irreducible. Using the inequalities

$$[K'(\varphi\alpha) : K'] \leqq [E' : K'] \leqq [E : K]$$

we now conclude that we must have an equality everywhere, and that $E' = K'(\varphi\alpha)$. This proves our proposition.

Next, we give the formalism for unramified extensions under various operations.

Proposition 11. *Let E be a finite extension of K.*

(i) *If $E \supset F \supset K$, then E is unramified over K if and only if E is unramified over F and F unramified over K.*

(ii) *If E is unramified over K, and K_1 is a finite extension of K, then EK_1 is unramified over K_1.*

(iii) *If E_1 and E_2 are finite unramified over K, then so is E_1E_2.*

Proof: The first assertion comes from the inequalities

$$[E' : K'] \leqq [E : K], \qquad [E' : F'] \leqq [E : F], \qquad [F' : K'] \leqq [F : K]$$

together with their multiplicativity property in towers. One must also use the fact that assertion (i) holds when "unramified" is replaced by "a finite separable extension". The second assertion is an immediate consequence of our criterion in Proposition 10. The third comes formally from the first and second.

Let us denote by φE the image under φ of the valuation ring $\mathfrak{o}_E$, for any finite extension E of K.

Proposition 12. *The map $E \to \varphi E$ gives a bijection of the finite unramified extensions of K and the finite separable extensions of $K' = \varphi K$.*

Proof: We have shown in Proposition 10 that every finite separable extension of K' is obtainable as an image φE, where E is unramified. We now must show uniqueness. If $E_1 \subset E_2$ are unramified, then clearly $\varphi E_1 \subset \varphi E_2$. It will therefore suffice to prove that if E_1, E_2 are two unramified extensions of K then $\varphi(E_1E_2) = \varphi E_1 \cdot \varphi E_2$. To do this, we can write $E_1 = K(\alpha_1)$ and $E_2 = K(\alpha_2)$ where α_1 and α_2 respectively satisfy the properties expressed in Proposition 10. Then $E_1E_2 = K(\alpha_1, \alpha_2)$, and we observe that with respect to the intermediate field E_1, α_2 satisfies a polynomial $f(X)$ such that f^{φ} has no multiple root. Hence $(E_1E_2)' = E_1'(\varphi\alpha_1) = K'(\varphi\alpha_1, \varphi\alpha_2)$ which proves Proposition 12,

§ 4. *Finite extensions*

Throughout this section, we shall deal with a field K having a nontrivial absolute value v. We wish to describe how this absolute value extends to finite extensions of K.

If we let K_v be the completion, we know that v can be extended to K_v, and then uniquely to its algebraic closure $\bar{K}_v$. If E is a finite extension of K, or even an algebraic one, then we can extend v to E by embedding E in $\bar{K}_v$ by an isomorphism over K, and taking the induced absolute value on E. We shall now prove that every extension of v can be obtained in this manner.

Proposition 13. *Let E be a finite extension of K. Let w be an absolute value on E extending v, and let E_w be the completion. Let K_w be the closure of K in E_w and identify E in E_w. Then $E_w = EK_w$ (the composite field).*

Proof: We observe that K_w is a completion of K, and that the composite field EK_w is algebraic over K_w and therefore complete by § 2, Proposition 3. Since it contains E, it follows that E is dense in it, and hence that $E_w = EK_w$.

If we start with an embedding $\sigma: E \to \bar{K}_v$ (always assumed to be over K), then we know again by § 2, Proposition 3 that $\sigma E \cdot K_v$ is complete. Thus this construction and the construction of the proposition are essentially the same, up to an isomorphism. In the future, we take the embedding point of view. We must now determine when two embeddings give us the same absolute value on E.

Given two embeddings $\sigma, \tau: E \to \bar{K}_v$, we shall say that they are *conjugate over* K_v if there exists an automorphism λ of $\bar{K}_v$ over K_v such that $\sigma = \lambda\tau$. We see that actually λ is determined by its effect on τE, or $\tau E \cdot K_v$.

PROPOSITION 14. *Let E be an algebraic extension of K. Two embeddings $\sigma, \tau: E \to \bar{K}_v$ give rise to the same absolute value on E if and only if they are conjugate over K_v.*

Proof: Suppose they are conjugate over K_v. Then the uniqueness of the extension of the absolute value from K_v to $\bar{K}_v$ guarantees that the induced absolute values on E are equal. Conversely, suppose this is the case. Let $\lambda: \tau E \to \sigma E$ be an isomorphism over K. We shall prove that λ extends to an isomorphism of $\tau E \cdot K_v$ onto $\sigma E \cdot K_v$ over K_v. Since τE is dense in $\tau E \cdot K_v$, an element $x \in \tau E \cdot K_v$ can be written

$$x = \lim \tau x_n$$

with $x_n \in E$. Since the absolute values induced by σ and τ on E coincide, it follows that the sequence $\lambda\tau x_n = \sigma x_n$ converges to an element of $\sigma E \cdot K_v$ which we denote by λx. One then verifies immediately that λx is independent of the particular sequence τx_n used, and that the map $\lambda: \tau E \cdot K_v \to \sigma E \cdot K_v$ is an isomorphism, which clearly leaves K_v fixed. This proves our proposition.

In view of the previous two propositions, if w is an extension of v to a finite extension E of K, then we may identity E_w and a composite extension EK_v of E and K_v. If $N = [E:K]$ is finite, then we shall call

$$N_w = [E_w: K_v]$$

the *local degree.* It is clear that $\sum_{w|v} [E_w: K_v] \leqq [E:K]$.

PROPOSITION 15. *Let E be a finite separable extension of K, of degree N. Then*

$$N = \sum_{w|v} N_w \,.$$

Proof. We can write $E = K(\alpha)$ for a single element α. Let $f(X)$ be its irreducible polynomial over K. Then over K_v, we have a decomposition

$$f(x) = f_1(x) \cdots f_r(X)$$

into irreducible factors $f_i(X)$. They all appear with multiplicity 1 according to our hypothesis of separability. The embeddings of E into $\bar{K}_v$ correspond to the maps of α onto the roots of the f_i. Two embeddings are conjugate if and only if they map α onto roots of the same polynomial f_i. On the other hand, it is clear that the local degree in each case is precisely the degree of f_i. This proves our proposition.

There exist interesting cases in characteristic p for which the relation in the proposition holds without the hypothesis of separability. This is notably the case for the discrete valuation arising from a simple subvariety of codimension 1 on an algebraic variety. For a sufficient condition under which the relation holds, including the special case just mentioned, the reader may see [3].

Whenever v is an absolute value on K such that for any finite extension E of K we have

$$[E:K] = \sum_{w|v} [E_w: K_v]$$

we shall say that v is *well behaved*. Suppose we have a tower of finite extensions, $L \supset E \supset K$. Let w range over the absolute values of E extending v, and u over those of L extending v. If $u \mid w$ then L_u contains E_w. Thus we have:

$$\begin{aligned} \sum_{u|v} [L_u: K_v] &= \sum_{w|v} \sum_{u|w} [L_u: E_w][E_w: K_v] \\ &= \sum_{w|v} [E_w: K_v] \sum_{u|w} [L_u: E_w] \\ &\leqq \sum_{w|v} [E_w: K_v][L:E] \\ &\leqq [E: K][L:E] \,. \end{aligned}$$

From this we immediately see that if v is well behaved, E finite over K, and w extends v on E, then w is well behaved (we must have an equality everywhere).

Let E be a finite extension of K. Let p^r be its inseparable degree. We recall that the norm of an element $\alpha \in K$ is given by the formula

$$N_K^E(\alpha) = \prod_{\sigma} \sigma\alpha^{p^r}$$

where σ ranges over all distinct isomorphisms of E over K (into a given algebraic closure).

If w is an absolute value extending v on E, then the norm from E_w to K_v will be called the *local norm.*

Replacing the above product by a sum, we get the trace, and the local trace. We abbreviate the trace by Tr.

PROPOSITION 16. *Let E be a finite extension of K, and assume that v is well behaved. Let $\alpha \in E$. Then:*

$$N_K^E(\alpha) = \prod_{w|v} N_{K_v}^{E_w}(\alpha)$$

$$Tr_K^E(\alpha) = \sum_{w|v} Tr_{K_v}^{E_w}(\alpha)$$

Proof: Suppose first that $E = K(\alpha)$, and let $f(X)$ be the irreducible polynomial of α over K. If we factor $f(X)$ into irreducible terms over K_v, then

$$f(X) = f_1(X) \cdots f_r(X)$$

where each $f_i(X)$ is irreducible, and the f_i are distinct because of our hypothesis that v is well behaved. The norm $N_K^E(\alpha)$ is equal to $(-1)^{\deg f}$ times the constant term of f, and similarly for each f_i. Since the constant term of f is equal to the product of the constant terms of the f_i, we get the first part of the proposition. The statement for the trace follows by looking at the penultimate coefficient of f and each f_i.

If E is not equal to $K(\alpha)$, then we simply use the transitivity of the norm and trace. We leave the details to the reader.

Taking into account Proposition 4 of § 2, we have:

THEOREM 2. *Let K have a well behaved absolute value v. Let E be a finite extension of K, and $\alpha \in E$. Let*

$$N_w = [E_w : K_v]$$

for each absolute w on E extending v. Then

$$\prod_{w|v} |\alpha|_w^{N_w} = |N_K^E(\alpha)|_v .$$

We conclude this section with some remarks on valuations which show how the ramification index and residue class degree remain the same when we pass to the completions. This will be especially useful in our applications to discrete valuations, taking into account Proposition 5 of § 2.

Suppose that we start with a field K with a non-trivial valuation v, corresponding valuation ring $\mathfrak{o}$, and maximal ideal $\mathfrak{m}$. Let us denote for a moment by $\hat{K}$ the completion of K, and let $\hat{\mathfrak{o}}$ resp. $\hat{\mathfrak{m}}$ be the closure of $\mathfrak{o}$ resp. $\mathfrak{m}$ in $\hat{K}$. Using the fact that if $|x| \neq |y|$ then $|x+y| = \max(|x|, |y|)$ we see that $\hat{\mathfrak{o}}$ is a valuation ring of $\hat{K}$, that $\hat{\mathfrak{m}}$ is its maximal ideal, and that

$$\hat{\mathfrak{o}} \cap K = \mathfrak{o}, \qquad \hat{\mathfrak{m}} \cap K = \mathfrak{m}.$$

Thus we have a canonical isomorphism

$$\mathfrak{o}/\mathfrak{m} \to \hat{\mathfrak{o}}/\hat{\mathfrak{m}}.$$

If E is a finite extension of K, and $\mathfrak{o}_E, \mathfrak{m}_E$ a valuation ring and its maximal ideal lying above $\mathfrak{o}, \mathfrak{m}$ and if we denote by a ^ the completion, then we have a canonical commutative diagram:

$$\begin{tikzcd}
\mathfrak{o}_E/\mathfrak{m}_E \arrow[r, "\approx"] & \hat{\mathfrak{o}}_E/\hat{\mathfrak{m}}_E \\
\mathfrak{o}/\mathfrak{m} \arrow[r, "\approx"] \arrow[u] & \hat{\mathfrak{o}}/\hat{\mathfrak{m}} \arrow[u]
\end{tikzcd}$$

the vertical arrows being inclusions. Thus the residue class field extension can be studied either over the given field K or its completion.

Similarly for the ramification index: If $x \in \hat{K}$ and $y \in K$ is such that $|y - x| < |x|$, then $|y - x + x| = |y|$ and hence the inclusion map

$$v(K^*) \to v(\hat{K}^*)$$

is an isomorphism (the value groups are the same). If now E is a finite extension of K, then we have the commutative diagram

$$\begin{tikzcd}
w(E^*) \arrow[r, "\approx"] & w(\hat{E}^*) \\
v(K^*) \arrow[r, "\approx"] \arrow[u] & v(\hat{K}^*) \arrow[u]
\end{tikzcd}$$

for any valuation w extending v to E, the vertical arrows being inclusions. Thus the ramification index can be studied locally over the completion.

Proposition 17. *Let v be a discrete, well behaved valuation on K, and let E be a finite extension of K. For each extension w of v to E, let*

$$e_w = (w(E^*) : v(K^*))$$

$$d_w = [\mathfrak{o}_w/\mathfrak{m}_w : \mathfrak{o}_v/\mathfrak{m}_v]$$

be the ramification index and residue class degree respectively. Then

$$\sum_{w|v} e_w d_w = [E : K]\ .$$

If E *is Galois over* K*, then all* e_w *are equal to the same number* e*, all* d_w *are equal to the same number* d*, and so*

$$[E:K] = edr$$

where r *is the number of extensions of* v *to* E*.*

Proof: Our first assertion comes from our assumption, and Proposition 5 of § 2. If E is Galois over K, we know that given two extensions w_1, w_2 of v there exists an automorphism σ of E over K such that $w_1 = w_2\sigma$. This symmetry implies our second assertion.

Historical Note:

The theory of absolute values is perfectly classical. We have, of course, not been exhaustive, giving only what we need most in the sequel. The reader wanting a more complete treatment may consult Bourbaki's *Commutative Algebra*, or Artin's *Algebraic Numbers and Algebraic Functions*. Actually, our treatment is much influenced by Artin's. Furthermore, although many special cases of the approximation theorem were known before (Chinese remainder theorems), it was first stated and proved in full generality for absolute values by Artin-Whaples [2].

CHAPTER II

The Product Formula

We shall give at the end of §1 the three most important examples of the type of field which concerns us. They are essentially fields of finite type over the prime field or over some constant field, the latter giving rise to relative theories.

On these fields, when we take suitable models of them, we get families of absolute values. These absolute values are either discrete or archimedean. A proper set of absolute values will be defined so as to axiomatize what is common to all of them, and to give us what we shall need later for certain applications, for instance, our study of height functions, or divisors and divisor classes.

It is particularly important when our set of absolute values will satisfy a product formula, the classical case being that of number fields. We place special emphasis on this case, and shall go deeper into the study of the group of ideal classes and the unit group, giving the classical results, which are more than qualitative. (In the case of units, they give the number of generators.)

In considering the examples of finite type (Examples 2 and 3) we limit ourselves to qualitative results, to the effect that certain groups of units and ideal classes are of finite type also. The proofs are essentially elementary, except for one reference to the Mordell-Weil theorem of Chapter V.

§1. *Proper sets of absolute values*

Let K be a field. An absolute value v on K is said to be *proper* if it is non-trivial, well behaved, and if, K having characteristic 0, its restriction to $\mathbf{Q}$ is either trivial, the ordinary absolute value, or a p-adic absolute value v_p.

A set M_K of absolute values on K is said to be *proper* if every absolute value in it is proper, if any two distinct absolute values are independent, and if, given $x \in K$, $x \neq 0$, there exists only a finite number of $v \in M_K$ such that $|x|_v \neq 1$. In particular, if M_K is proper, there can be only a finite number of archimedean absolute values in M_K.

If E is an algebraic extension of K, we shall denote by M_E the set of absolute values on E extending some absolute value in M_K. If E is finite over K, then M_E is proper if M_K is proper. (Its absolute values are well behaved by Chapter I, §4, and the other conditions are trivially verified.)

Let M_K be a proper set of absolute values on K. For each $v \in M_K$, let λ_v be a real number >0. We shall say that M_K satisfies the *product formula with multiplicities* λ_v if for each $x \in K$, $x \neq 0$, we have

$$\prod_{v \in M_K} |x|_v^{\lambda_v} = 1 .$$

By assumption, there is only a finite number of terms in this product which are not equal to 1, so the product makes sense. We shall say that M_K satisfies the *product formula* if all $\lambda_v = 1$. When we deal with a fixed set of multiplicities λ_v, then we write for convenience

$$\| x \|_v = |x|_v^{\lambda_v}$$

so that our product formula reads

$$\prod_{v \in M_K} \| x \|_v = 1 .$$

Suppose now that we have a field $\mathbf{F}$ with a proper set $M_\mathbf{F}$ of absolute values satisfying the product formula with multiplicities 1. Let K be a finite extension of $\mathbf{F}$, and let M_K be the set of absolute values on K which extend the absolute values of $M_\mathbf{F}$. Then M_K is also a proper set of absolute values on K. If $v_0 \in M_\mathbf{F}$ and $v \in M_K$, with $v \mid v_0$, set $N_v = [K_v : \mathbf{F}_{v_0}]$. Then for any element $\alpha \in E$, $\alpha \neq 0$, we get by Theorem 2 of Chapter I, §4

$$\begin{aligned} 1 &= \prod_{v_0 \in M_\mathbf{F}} | N_F^K(\alpha) |_{v_0} \\ &= \prod_{v_0 \in M_\mathbf{F}} \prod_{v \mid v_0} |\alpha|_v^{N_v} \\ &= \prod_{v \in M_K} |\alpha|_v^{N_v} . \end{aligned}$$

This shows that M_K satisfies the product formula with multiplicities N_v.

Example 1

The classical example is that of the rational numbers $\mathbf{Q}$. For each prime number p we have the absolute value v_p described in Chapter I, § 1. The ordinary archimedean absolute value will be said to belong to the *prime at infinity*, and we shall also use p or p_∞ to denote its prime. If l is a prime number, then

$$|\,l\,|_p = \begin{cases} 1 & \text{if } p \text{ is a prime number } \neq l \\ 1/l & \text{if } p \text{ is a prime number } = l \\ l & \text{if } p \text{ is the prime at infinity .} \end{cases}$$

Thus the set $M_{\mathbf{Q}}$ of all absolute values v_p (including the one at infinity) satisfies the product formula, the above argument showing this for prime numbers of $\mathbf{Z}$, the general case following by multiplicativity. It is also clear that our set is proper (using Proposition 15 of Chapter I, § 4).

If K is a number field, then the set M_K of absolute values extending those of $M_{\mathbf{Q}}$ will be called the *canonical set*.

If $I = I_K$ is the ring of integers of K (i.e. the integral closure of $\mathbf{Z}$ in K) then the valuation rings of K are in bijective correspondence with the prime ideals $\mathfrak{p}$ of I. We shall always exclude the trivial one. Each valuation ring is of type $I_{\mathfrak{p}}$, the local ring of I at $\mathfrak{p}$ (*IAG*, Theorem 4 of Chapter I, § 6). The primes of K corresponding to these valuation rings will be called the *finite primes* of K. Those primes corresponding to the absolute values extending the ordinary absolute value on $\mathbf{Q}$ will be called the *infinite primes*, or *primes at infinity*.

Let K be a number field, and v one of the absolute values extending the ordinary absolute value on $\mathbf{Q}$. Then K_v is either the field of real numbers of the field of complex numbers. We then say that v is *real* or *complex*, accordingly, and use a similar terminology for its associated prime. The multiplicity

$$N_v = [K_v : \mathbf{Q}_v]$$

is then 1 or 2 according as v is real or complex, and thus for $x \in K$, we have

$$\|\,x\,\|_v = |\,x\,|_v^{N_v}$$

either $|\,x\,|$ or $|\,x\,|^2$.

Let us now consider a finite prime corresponding to a prime ideal $\mathfrak{p}$ in the ring of integers I of K. Then there is a unique prime number p lying in $\mathfrak{p}$ and we denote by $v_{\mathfrak{p}}$ or $|\ |_{\mathfrak{p}}$ the absolute beloning to $\mathfrak{p}$ extending the p-adic absolute value on $\mathbf{Q}$. Its multiplicity is

$$N_{\mathfrak{p}} = [K_{\mathfrak{p}} : \mathbf{Q}_p] = e_{\mathfrak{p}} d_{\mathfrak{p}}$$

where $e_{\mathfrak{p}}$ and $d_{\mathfrak{p}}$ are the ramification index and residue class degree, respectively. Letting $\mathfrak{o}_{\mathfrak{p}}$ and $\mathfrak{m}_{\mathfrak{p}}$ be the valuation ring and maximal ideal corresponding to $\mathfrak{p}$, we see that the number of elements in $\mathfrak{o}_{\mathfrak{p}}/\mathfrak{m}_{\mathfrak{p}}$, which we denote by $\mathbf{N}\mathfrak{p}$, is

$$\mathbf{N}\mathfrak{p} = p^{d_{\mathfrak{p}}} ,$$

since $\mathfrak{o}_{\mathfrak{p}}/\mathfrak{m}_{\mathfrak{p}}$ is of degree $d_{\mathfrak{p}}$ over $\mathbf{Z}/p\mathbf{Z}$. Furthermore, let π be a prime element at $\mathfrak{p}$. Then by definition,

$$| \pi^{e_{\mathfrak{p}}} |_{\mathfrak{p}} = | p |_{\mathfrak{p}} = 1/p .$$

Hence

$$| \pi |_{\mathfrak{p}}^{e_{\mathfrak{p}} d_{\mathfrak{p}}} = (1/\mathbf{N}\mathfrak{p}) .$$

From this we get:

Proposition 1. *Let K be a number field, $\mathfrak{p}$ a finite prime, $\mathbf{N}\mathfrak{p}$ the order of $\mathfrak{o}_{\mathfrak{r}}/\mathfrak{m}_{\mathfrak{p}}$, and $| \ |_{\mathfrak{p}}$ the absolute value extending the p-adic absolute value on $\mathbf{Q}$. Let $|| \ ||_{\mathfrak{p}} = | \ |_{\mathfrak{p}}^{N_{\mathfrak{p}}}$. Then for $x \in K$, $x \neq 0$,*

$$|| x ||_{\mathfrak{p}} = (1/\mathbf{N}\mathfrak{p})^{\mathrm{ord}_{\mathfrak{p}} x} .$$

As our set of absolute values in M_K is in bijective correspondence with the primes of K, we write our product formula in the form

$$\prod_{\mathfrak{p} \in M_K} || x ||_{\mathfrak{p}} = 1 .$$

Example 2

Aside from algebraic number fields, the most important example of a proper set of absolute values satisfying the product formula is that obtained from the subvarieties of codimension 1 on a projective variety, non-singular in codimension 1. We shall now describe this situation.

Let k be a field, V a projective variety defined over k, and non-singular in codimension 1. A *prime rational divisor* $\mathfrak{p}$ of V over k is a cycle of codimension 1,

$$\mathfrak{p} = \mu \sum_{\sigma} W^{\sigma}$$

where W is subvariety of V, of codimension 1, defined over an algebraic extension of k, W^σ ranges over the distinct conjugates of W over k, and μ is its order of inseparability over k. Such a prime rational divisor (also called a prime divisor) determines its local ring in the function field $k(V)$ of V over k, and this local ring is a discrete valuation ring. Thus we have the notion order at $\mathfrak{p}$ of a function $x \in k(V)$, $x \neq 0$. By the *degree* of $\mathfrak{p}$, noted deg $\mathfrak{p}$, we shall mean the projective degree, i.e. the number of points of intersection with a generic linear variety of complementary dimension in the given projective embedding. Similarly, we have the degree of an arbitrary divisor, i.e. the number of points of intersection with a linear variety, generic over a field of definition for all its components. If $x \in k(V)$, $x \neq 0$, and (x), $(x)_0$, $(x)_\infty$ denote the divisor of x, its divisor of zeros, and its polar divisor respectively, then

$$\deg (x)_0 = \deg (x)_\infty$$

and hence $\deg (x) = 0$.

Now let c be a number, $0 < c < 1$, and for each prime divisor $\mathfrak{p}$ of V over k, let, for $x \neq 0$,

$$v_\mathfrak{p}(x) = | x |_\mathfrak{p} = c^{(\mathrm{ord}_\mathfrak{p} x) \deg (\mathfrak{p})} .$$

For $K = k(V)$, we then get a family $M_K = M_{k(V)}$ of proper absolute values, which satisfy the product formula in view of the remark we just made. This will be said to be the family associated with the model V of the field K. Once our number c is chosen, we have a bijective correspondence between prime divisiors and absolute values in $M_{k(V)}$, and thus we shall also write our product formula

$$\prod_{\mathfrak{p} \in M_{k(V)}} | x |_\mathfrak{p} = 1 .$$

As we have multiplicities $=1$, we have $\| x \|_\mathfrak{p} = | x |_\mathfrak{p}$.

By a *function field* K over the constant field k we shall always mean a finitely generated, regular extension of k. Each model V of K over k which is projective and non-singular in codimension 1 gives rise to absolute values as above. This slightly restricts the notion of function fields as used in the classical theory of function fields in one variable, i.e. finitely generated extensions K of a field k which are of transcendence degree 1 and such K is algebraically closed in k. For such fields, one also has a product formula for the set of discrete valuations, trivial on k, such valuations being well behaved and hence proper.

Example 3

Let R be a Noetherian ring, which we assume to be integrally closed. If $\mathfrak{p}$ is a minimal prime ideal of R not equal to 0, then the local ring $R_{\mathfrak{p}}$ is a discrete valuation ring in the quotient field of R, and an element of R has only a finite number of zeros from among such $\mathfrak{p}$. Suitably normalizing the absolute values we get from such a $\mathfrak{p}$, in case they induce a non-trivial absolute value on the prime ring, we get a proper set of absolute values on the quotient field of R, which will be denoted by M_R, or $M_{R,k}$.

By a *variety scheme* over a ring R, we shall mean a scheme V over spec (R) which is irreducible, of finite type, having no divisors of 0 in its local rings, and such that if k is the field of fractions of R, and V^k the extension of V to the ground field k, then the field of rational functions on V^k is a regular extension of k. Then V^k is essentially the same thing as a variety, defined over k. (For schemes, see [11].)

A variety scheme V over R will be said to be *normal* if each one of its local rings is integrally closed, and if R is integrally closed.

Now let R be Noetherian, and V a variety scheme over R. Assume that R is integrally closed, and V normal. Let k be the field of fractions of R, and K the function field of V, so that K contains k.

We then get, on the one hand, a family of discrete valuations on K which are trivial on k, corresponding to the prime divisors on V^k over k. (If V is affine, a prime divisor in this case is determined by a minimal non-zero prime ideal of the affine ring of V^k over k. It need not be absolutely simple, merely k-simple.) This set will be denoted by $M_{k,V}$.

We get however other discrete valuations. For each prime ideal $\mathfrak{p}$ of R which is minimal not equal to 0, the local ring $R_{\mathfrak{p}}$ is, as we have already pointed out, a discrete valuation ring of k. The fiber of V above the point $\mathfrak{p}$ in spec(R) is either empty, or consists of a finite number of components, each of which gives rise to a discrete valuation on K, corresponding to what we shall call a *prime divisor on* V *lying above* $\mathfrak{p}$. We denote by $M_{R,V}$ the set of such discrete valuations.

We let M_V be the union of $M_{R,V}$ and $M_{k,V}$. Then M_V is a set of discrete valuations on K, and each non-zero element of K is a unit at all but a finite number of them. By suitably normalizing the absolute values corresponding to these valutions which induce non-trivial valuations on $\mathbf{Q}$, we get a proper set of absolute values on K, also denoted by M_V. It does not in general satisfy the product formula.

We recall that in view of our assumptions on V, and R, all but a finite number or primes $\mathfrak{p} \in M_R$ are such that there is exactly one prime of $M_{R,V}$ lying above it, corresponding to a "non-degenerate" reduction of V mod $\mathfrak{p}$. We shall denote by $M'_{R,V}$ the subset of $M_{R V}$ lying above such $\mathfrak{p}$ in M_R, so that $M'_{R,V}$ can be identified with a subset M'_R of M_R, differing from M_R by only a finite set of elements.

In studying variety schemes above a ring, the case in which all objects are of finite type over the prime ring is particularly important, and we shall frequently use the following result.

Proposition 2. *Let R be a ring finitely generated over the prime ring. Then there exists an element $a \neq 0$ in the prime ring such that the integral closure of $R[1/a]$ (in its quotient field) is a finite module over $R[1/a]$.*

Proof. The case of characteristic >0 is standard, since the prime ring is a field (Noether's Normalization Theorem). The case of interest to us here is that of characteristic 0. Denote by $R_{\mathbf{Q}}$ the algebra generated over $\mathbf{Q}$ by R. By the normalization theorem, there exist elements $t_1, \cdots, t_r \in R$ algebraically independent over $\mathbf{Q}$ such that $R_{\mathbf{Q}}$ is integral over $\mathbf{Q}[t_1, \cdots, t_r]$. Since R is finitely generated over $\mathbf{Z}$, there exists an element $a \in \mathbf{Z}$, $a \neq 0$ such that R is integral over $\mathbf{Z}[1/a, t_1, \cdots, t_r]$. Since this latter ring is integrally closed, its integral closure in the quotient field of R is a finite module over it, say by IAG, Theorem 2 of Chapter V, § 1.

Remark. It is in fact true that the integral closure of R itself is a finite module over R, but this theorem, due to Nagata, is very much more difficult to prove.

§ 2. *Divisors, ideals, and units*

Let K be a field with a proper set of absolute values M_K. By an *M_K-divisor* one means a real valued function

$$v \to \mathfrak{d}(v)$$

of the absolute values in M_K with the following properties:

(i) $\mathfrak{d}(v) > 0$ for all $v \in M_K$, and $\mathfrak{d}(v) = 1$ for all but a finite number of the v's.

(ii) For each non-archimedean $v \in M_K$, there exists $\alpha \in K^*$ such that $\mathfrak{d}(v) = v(\alpha)$.

If $\alpha \in K^*$ then the function $v(\alpha)$ on M_K is an M_K-divisor, denoted by $\mathfrak{d}_\alpha$, and is called *principal.*

We observe that the M_K-divisors are closed under the operations of sup and inf, and that they form a group under multiplication (componentwise).

If E is an algebraic extension of K, then for each $w \in M_E$ we extend an M_K-divisor $\mathfrak{d}$ to M_E by the formula

$$\mathfrak{d}(w) = \mathfrak{d}(v)$$

if w extends $v \in M_K$. If E is finite over K, then this is an M_E-divisor, and our extension maps the M_K-divisors injectively into the M_E-divisors. By abuse of language, even if E is infinite, we sometimes say that our extension is an M_E-divisor (although condition (i) might be violated).

Suppose that v is discrete. If π is a prime element at v, then

$$\mathfrak{d}(v) = |\pi|_v^n$$

for some integer n, which is called the *order* of $\mathfrak{d}$ at v, and abbreviated $\operatorname{ord}_v\mathfrak{d}$, or $\operatorname{ord}_\mathfrak{p}\mathfrak{d}$ if $\mathfrak{p}$ is the prime associated with v.

Let us denote by M_K^0 the subset of M_K consisting of the non-archimedean absolute values.

Suppose that all the absolute values M_K^0 are discrete, and that for each M_K-divisor we take its restriction to these non-archimedean absolute values. These restrictions will be called *finite divisors*, and they form a group which is clearly isomorphic in the natural manner with the free abelian group generated by the primes corresponding to the absolute values in M_K^0. This latter group will be called the group of *M_K-ideals.* Each element $x \in K$, $x \neq 0$ determines an M_K-ideal denoted by (x), and called a *principal M_K-ideal.* The principal M_K-ideals form a group, and the factor group is called the group of *M_K-ideal classes.*

If $\mathfrak{a}$ is an M_K-ideal, then we can write

$$\mathfrak{a} = \sum_{\mathfrak{p}} n_\mathfrak{p}\mathfrak{p}$$

with $n_\mathfrak{p} \in \mathbf{Z}$, and $\mathfrak{p}$ ranging over the primes of M_K^0, all but a finite number of $n_\mathfrak{p}$ being 0. We write

$$n_{\mathfrak{p}} = \operatorname{ord}_{\mathfrak{p}} \mathfrak{a} .$$

If $\alpha \in K^*$, then $\operatorname{ord}_{\mathfrak{p}} \alpha = \operatorname{ord}_{\mathfrak{p}} (\alpha)$, i.e. it is the order of its divisor at $\mathfrak{p}$.

If two M_K-ideals $\mathfrak{a}$ and $\mathfrak{b}$ lie in the same ideal class, we write

$$\mathfrak{a} \sim \mathfrak{b} .$$

This is the notion of linear equivalence, in Example 2 of § 1. An M_K-ideal is said to be *integral* if $\operatorname{ord}_{\mathfrak{p}} \mathfrak{a} \geqq 0$ for all $\mathfrak{p}$ in M_K^0.

By an *M_K-unit* we shall mean an element $\alpha \in K^*$ such that

$$| \alpha |_v = 1$$

for all non-archimedean absolute values v of M_K. The M_K-units obviously form a multiplicative group, which we denote by U_{M_K}, U_K, or simply U.

In Example 1 of § 1, our group of M_K-ideals is isomorphic in the obvious manner with the classical group of ideals of the ring of integers (including the fractional ideals, of course). The M_K-units are simply the units of the number field in the classical sense.

In Example 2 of § 1, the group of M_K-ideals is simply the group of divisors on the variety, rational over the constant field k. The units are the non-zero constants (since we assume the variety complete and non-singular in codimension 1).

In Example 3, we have:

Proposition 3. *Let R be a Notherian, integrally closed ring. Let M_R be its set of primes as in Example 3 of § 1. Then R^* is the set of M_R-units of the field K of fractions of R.*

Proof. It is well known from commutative algebra that $R = \bigcap R_{\mathfrak{p}}$, the intersection being taken over all $\mathfrak{p} \in M_R$. If $u \in K$ is an M_R-unit, then u lies in each $R_{\mathfrak{p}}$ since, in particular, it has no pole among the $\mathfrak{p}$'s. Hence it is in R, and since it has no zero, it must be invertible (Krull's principal ideal theorem). The converse is clear.

Starting with a proper set of absolute values M_K on a field K, we shall frequently let S be a finite subset, containing the archimedean absolute values which may occur in M_K. It is customary to call *S-units* the $(M_K - S)$-units, i.e. those elements of K having absolute value 1 for all $v \notin S$. These form a multiplicative group K_S. We shall study in § 4 the qualitative effect of removing such a finite set from consideration. For number fields, we shall give in § 3 the number of generators of the S-units.

Let us now consider a number field K with its canonical set of absolute values M_K. For each finite prime, we had defined $\mathbf{N}\mathfrak{p}$ to be the number of elements in the residue class field of its valuation ring. We extend this to M_K-ideals by linearity, so that

$$\mathbf{N}\mathfrak{a} = \prod (\mathbf{N}\mathfrak{p})^{\mathrm{ord}_{\mathfrak{p}}\mathfrak{a}} .$$

It is clear that there is only a finite number of $\mathfrak{p}$'s such that $\mathbf{N}\mathfrak{p}$ lies below a given number, and hence that there is only a finite number of integral M_K-ideals $\mathfrak{a}$ such that $\mathbf{N}\mathfrak{a}$ lies below a given number.

We shall now prove that the M_K-ideal classes form a finite group. We need a lemma.

Lemma 1. *Let I be the ring of integers of the number field K. Let $\mathfrak{a}$ be an integral M_K-ideal, $\mathfrak{a} = \sum n_{\mathfrak{p}}\mathfrak{p}$. Then* $\mathbf{N}\mathfrak{a}$ *is the number of elements of the module*

$$\prod I_{\mathfrak{p}}/\mathfrak{m}_{\mathfrak{p}}^{n_{\mathfrak{p}}}$$

(*the product being taken over the finite primes, and all but a finite number of terms being trivial*).

Proof. The valuation ring $I_{\mathfrak{p}}$ is a discrete valuation ring, and its maximal ideal $\mathfrak{m}_{\mathfrak{p}}$, generated by π, is principal. As a module over $I_{\mathfrak{p}}/\mathfrak{m}_{\mathfrak{p}}$,

$$\pi^r I_{\mathfrak{p}}/\pi^{r+1} I_{\mathfrak{p}}$$

is isomorphic to $I_{\mathfrak{p}}/\mathfrak{m}_{\mathfrak{p}}$ itself, under multiplication by π^r, r being an integer $\geqq 0$. We have the chain of ideals

$$I_{\mathfrak{p}} \supset \pi I_{\mathfrak{p}} \supset \pi^2 I_{\mathfrak{p}} \supset \cdots \supset \pi^r I_{\mathfrak{p}}$$

and from this we see that the number of elements in the module $I_{\mathfrak{p}}/\pi^r I_{\mathfrak{p}}$ is equal to $(\mathbf{N}\mathfrak{p})^r$. By definition, this is equal to $\mathbf{N}(r\mathfrak{p})$ for the M_K-ideal $r\mathfrak{p}$. Our assertion is now obvious by linearity.

Proposition 4. *Let K be a number field, $\alpha \in K^*$, and $\mathfrak{a} = (\alpha)$ its M_K-ideal. Then for each prime number p, we have*

$$\mathrm{ord}_p \mathbf{N}\mathfrak{a} = \mathrm{ord}_p N_{\mathbf{Q}}^{K}(\alpha)$$

and consequently

$$|\mathbf{N}\mathfrak{a}| = |N_{\mathbf{Q}}^{K}(\alpha)| = \mathbf{N}\mathfrak{a}$$

where $|\ |$ *is the ordinary absolute value.*

Proof. For each $\mathfrak{p} \mid p$, let $d_{\mathfrak{p}}$ be the residue class degree. By Proposition 16 of Chapter I, § 4 we have

$$N^K_{\mathbf{Q}}(\alpha) = \prod_{\mathfrak{p} \mid p} N^{K_{\mathfrak{p}}}_{\mathbf{Q}_p}(\alpha)$$

and hence, using Proposition 6 of Chapter I, § 2,

$$\begin{aligned} \operatorname{ord}_p N^K_{\mathbf{Q}}(\alpha) &= \sum_{\mathfrak{p} \mid p} \operatorname{ord}_p N^{K_{\mathfrak{p}}}_{\mathbf{Q}_p}(\alpha) \\ &= \sum_{\mathfrak{p} \mid p} d_{\mathfrak{p}} \operatorname{ord}_{\mathfrak{p}} \alpha \,. \end{aligned}$$

Our first assertion now follows by definition, and the second is obvious from the first.

THEOREM 1. *Let K be a number field. Then there is a constant C depending on K only, such that for any integral M_K-ideal $\mathfrak{a}$ there exists an integral ideal $\mathfrak{b} \sim \mathfrak{a}$ such that $\mathbf{N}\mathfrak{b} \leqq C$, and the group of M_K-ideal classes is finite.*

Proof. Our second assertion is an immediate consequence of the first and of the fact that there are only a finite number of integral M_K-ideals $\mathfrak{b}$ with $\mathbf{N}\mathfrak{b} < C$. Let us prove our first assertion.

Let I be the integers of K. It is a finite module over $\mathbf{Z}$, which is a principal ideal ring, and thus has a basis $\omega_1, \cdots, \omega_N$ over $\mathbf{Z}$. Let S be the set of elements of I of type

$$a_1\omega_1 + \cdots + a_N\omega_N$$

with a_i integers, such that

$$0 \leqq a_i \leqq (\mathbf{N}\mathfrak{a})^{1/N} + 1 \,.$$

Then there are more than $\mathbf{N}\mathfrak{a}$ elements in S, and thus there are two distinct elements α, β in S such that $\alpha - \beta = \xi$ will map into 0 in the homomorphism

$$I \to \prod I_{\mathfrak{p}}/\mathfrak{m}_{\mathfrak{p}}^{\operatorname{ord}_{\mathfrak{p}} \mathfrak{a}}$$

(referring to Lemma 2). It follows from this that $(\xi) \geqq \mathfrak{a}$, and that we can write

$$(\xi) = \mathfrak{a} + \mathfrak{b}$$

with an M_K-integral ideal $\mathfrak{b}$. On the other hand, we estimate

$$|\, N^K_{\mathbf{Q}}(\xi) \,| = \prod_{\sigma} |\, c_1\omega_1^{\sigma} + \cdots + c_N\omega_N^{\sigma} \,|$$

where $0 \leqq c_i \leqq (\mathbf{N}\mathfrak{a})^{1/N} + 1$, and we see that there is a constant C (depending on the maximum of the archimedean absolute values of the ω_i, and on N) such that

$$| N^K_{\mathbf{Q}}(\xi) | \leqq C\mathbf{N}\mathfrak{a} .$$

Using Proposition 4, we get $\mathbf{N}\mathfrak{b} \leqq C$, and $\mathfrak{b} \sim \mathfrak{a}$ by definition. This proves our theorem.

§ 3. *Unit theorem*

Let K be a field with a proper family M_K of absolute values. For each M_K-divisor $\mathfrak{d}$ we shall denote by $L(\mathfrak{d})$ the set of elements $x \in K$ such that for each $v \in M_K$ we have

$$| x |_v \leqq \mathfrak{d}(v) .$$

If $\alpha \in K^*$, and if we denote by $\alpha\mathfrak{d}$ the product $\mathfrak{d}_\alpha \mathfrak{d}$ of the principal M_K-divisor α times $\mathfrak{d}$, then $L(\alpha\mathfrak{d})$ and $L(\mathfrak{d})$ are in canonical bijection under the mapping

$$x \to \alpha x \qquad x \in L(\mathfrak{d}) .$$

We denote the number of elements of $L(\mathfrak{d})$ by $\lambda(\mathfrak{d})$. Then $\lambda(\alpha\mathfrak{d}) = \lambda(\mathfrak{d})$. If we think of $\mathfrak{d}$ as prescribing the sides of a box, all but a finite number of which are 1, then $\lambda(\mathfrak{d})$ may be interpreted as the number of field elements in the box.

We define $| \mathfrak{d} |_v = \mathfrak{d}(v)$, and when we have multiplicities N_v,

$$|| \mathfrak{d} ||_v = \mathfrak{d}(v)^{N_v} .$$

We define the *K-size*, or *size* of $\mathfrak{d}$ to be

$$|| \mathfrak{d} ||_K = \prod_{v \in M_K} || \mathfrak{d} ||_v .$$

If M_K satisfies the product formula with mulitiplicities N_v, then the size of $\mathfrak{d}$ is the same as that of $\alpha\mathfrak{d}$. This size may be interpreted as giving the volume of our box. We shall now prove that in number fields, the number of elements in the box is approximately equal to the volume.

THEOREM 2. *Let K be a number field with its canonical proper set of absolute values M_K. There exist two numbers $c_1, c_2 > 0$ depending only on K, such that for any M_K-divisor $\mathfrak{d}$, we have*

$$c_1 || \mathfrak{d} ||_K < \lambda(\mathfrak{d}) \leqq \sup [1, c_2 || \mathfrak{d} ||_K] .$$

Proof. Suppose that for some archimedean absolute value v_0 in M_K, the completion K_{v_0} is complex. In this complex plane, consider the square centered at the origin, with sides of length $2\mathfrak{d}(v_0)$. Let m be an integer such that

$$m < \lambda(\mathfrak{d})^{1/2} \leqq m + 1 .$$

Without loss of generality, we may assume $m \neq 0$, and so $m \geqq 1$. Cut up each side of the square into m equal parts, thus giving rise to m^2 small squares inside the big one. Our set $L(\mathfrak{d})$ is embedded inside the big square at K_{v_0}. Since it contains more than m^2 elements, there exist two distinct elements $x, y \in L(\mathfrak{d})$ lying in the same small square. Hence for their difference, we get the estimate

$$|x - y|_{v_0} \leqq \frac{2\mathfrak{d}(v_0)}{m} .$$

If v is any other archimedean absolute value of M_K, we get

$$|x - y|_v \leqq 2\mathfrak{d}(v)$$

and if v is non-archimedan, then

$$|x - y|_v \leqq \mathfrak{d}(v) .$$

Taking the product, we get

$$1 = \prod_{v \in M_K} |x - y|_v^{N_v} = 4^s \|\mathfrak{d}\|_K / m^2$$

where s is the number of archimedean primes in M_K. Since $(m + 1)^2 \leqq 4m^2$ the inequality on the right follows immediately.

If there is no complex absolute value in M_K, then we proceed in a similar manner, using a real one v_0, and cut up the interval centered at the origin of length $2\mathfrak{d}(v_0)$ into m equal pieces, giving rise to m small intervals, with

$$m < \lambda(\mathfrak{d}) \leqq m + 1 .$$

The rest of the argument follows in the same way

Let us prove the other inequality. Let I_K be the integers of K (i.e. the integral closure of $\mathbf{Z}$ in K) and let $\omega_1, \cdots, \omega_N$ be a basis of I_K over $\mathbf{Z}$, so that $N = [K : \mathbf{Q}]$. Put

$$c_0 = N \sup_{v,i} |\omega_i|_v$$

the sup being taken over the archimedean absolute values v in M_K

and over i. This is a number depending only on K.

Let $\mathfrak{d}$ be our given M_K-divisor. By the approximation theorem (Theorem 1 of Chapter 1, § 1) there exists an element $\alpha \in K^*$ such that

$$c_0 \leqq |\alpha\mathfrak{d}|_v \leqq 2c_0$$

for each archimedean absolute value v in M_K. We now select an element $a \in \mathbf{Z}$, $a \neq 0$, such that $a\alpha\mathfrak{d}$ has absolute value $\leqq 1$ at all non-archimedean $v \in M_K$, just by taking a highly divisible at a lot of prime numbers. In view of the fact that $\lambda(\mathfrak{d})$ and $\|\mathfrak{d}\|_K$ do not change if we multiply $\mathfrak{d}$ by an element of K^*, we may therefore assume, without loss of generality, that our M_K-divisor satisfies the inequalities

$$c_0 |a|_v \leqq |\mathfrak{d}|_v \leqq 2c_0 |a|_v$$

for some element $a \in \mathbf{Z}$, $a > 0$.

We must exhibit elements of $L(\mathfrak{d})$. For this purpose, consider the set L of elements of I_K consisting of those which can be expressed in the form

$$a_1\omega_1 + \cdots + a_N\omega_N$$

with $a_i \in \mathbf{Z}$, and $0 \leqq a_i \leqq a$. Then our set L contains more than a^N elements.

For each non-archimedean prime $\mathfrak{p}$ in M_K, let $\nu_{\mathfrak{p}} = \operatorname{ord}_{\mathfrak{p}}\mathfrak{d}$. Let $\mathfrak{o}_{\mathfrak{p}}, \mathfrak{m}_{\mathfrak{p}}$ be the discrete valuation ring and its maximal ideal, and consider the additive group

$$\prod \mathfrak{o}_{\mathfrak{p}}/\mathfrak{m}_{\mathfrak{p}}^{\nu_{\mathfrak{p}}}$$

which has $\Pi(\mathbf{N}\mathfrak{p})^{\nu_{\mathfrak{p}}}$ elements. In the additive homomorphism of I_K into this group, we look at the image of L. There will be a subset L' of L with at least

$$\frac{a^N}{\prod(\mathbf{N}\mathfrak{p})^{\nu_{\mathfrak{p}}}}$$

elements, all of which have the same image. Take one fixed element $x \in L'$, and let y range over L'. Then for each non-archimedean absolute value v in M_K we have

$$|x - y|_v \leqq \mathfrak{d}(v)$$

because $\operatorname{ord}_{\mathfrak{p}}(x - y) \geqq \operatorname{ord}_{\mathfrak{p}} \mathfrak{d}$. If v is archimedean, then

$$|x-y|_v \leqq c_0 |a|_v \leqq \mathfrak{d}(v)$$

by an obvious estimate. Thus our elements $x - y$ lie in $L(\mathfrak{d})$. We have therefore proved that

$$\lambda(\mathfrak{d}) \geqq a^N \prod \frac{1}{\mathbf{N}\mathfrak{p}^{\nu_\mathfrak{p}}} .$$

We observe finally that

$$a^N = \prod_{v | v_\infty} |a|_v^{N_v} > c_1 \prod_{v | v_\infty} |\mathfrak{d}|_v^{N_v}$$

the product being taken over the archimedean absolute values, and c_1 being an obvious constant, while

$$1/\mathbf{N}\mathfrak{p}^{\nu}\mathfrak{p} = ||\mathfrak{d}||_v$$

if v is the non-archimedean absolute value belonging to $\mathfrak{p}$. Taking the full product over all absolute values proves our other inequality.

Let K be a number field, and S a finite subset of M_K containing the archimedean absolute values. Let s be the number of elements of S. Let K_S be the S-units, i.e. the elements $x \in K^*$ such that

$$||x||_v = 1 \qquad v \notin S .$$

We map K_S into Euclidean s-space as follows. Let $v_1, \cdots, v_s$ be the absolute values of S. Map

$$x \to (\log ||x||_1, \cdots, \log ||x||_s) ,$$

and call this map

$$\log\colon K_S \to \mathbf{R}^s .$$

By the product formula, the image of K_S is contained in the hyperplane defined by the equation

$$\xi_1 + \cdots + \xi_s = 0 ,$$

so that this image is at most $(s-1)$-dimensional. The following theorem describes this image.

THEOREM 3. *Let K be a number field, S a finite subset of M_K containing the archimedean absolute values, and s the number of elements of S. Let K_S be the S-units. Then $\log(K_S)$ is an $(s-1)$-dimensional lattice.*

By saying that it is a lattice, we mean that it is a discrete subgroup of $\mathbf{R}^s$, and by saying that it is $(s-1)$-dimensional, we mean that the vector space generated by it is the entire hyperplane mentioned above. Thus in particular, it follows that $\log(K_S)$ is a free abelian group on $s-1$ generators. Since the kernel of our log mapping is clearly the set of roots of unity in K (its elements form a group, and have bounded absolute value, so form a finite group), we get as a corollary:

COROLLARY. *Let K be a number field, S a finite subset of M_K containing the archimedean absolute values. Then K_S modulo the group of roots of unity in K is a free abelian group on s generators ($s =$ number of elements of S).*

Let us now prove the theorem, which is known as the *unit theorem.*

We begin by observing that in any bounded region of $\mathbf{R}^s$ there exists only a finite number of elements of $\log(K_S)$. Indeed, if $\log(x)$ lies in such a region, then the absolute values of x and its conjugates must be bounded, and hence x can satisfy only a finite number of equations of degree $\leqq [K:\mathbf{Q}]$ over $\mathbf{Q}$, because the coefficients of such equations are elementary symmetric functions of x and its conjugates. By a well known property of Euclidean space, whose proof we shall recall at the very end, it follows that $\log(K_S)$ is a discrete, finitely generated subgroup of $\mathbf{R}^s$. We must now show that it has dimension $s-1$.

For this purpose, we shall first prove that given an index i, there exists a vector $(\xi_1, \cdots, \xi_s)$ in $\log(K_S)$ such that $\xi_i > 0$ and $\xi_j < 0$ for $j \neq i$. We shall then prove that any $s-1$ such vectors are linearly independent over $\mathbf{R}$.

We need the following lemma:

LEMMA 2. *Given $v_0 \in M_K$ there exists a number $c(v_0) > 0$ such that for any M_K-divisor $\mathfrak{d}$ there exists $\beta \in K^*$ such that*

$$1 \leqq \|\beta\mathfrak{d}\|_v \leqq c(v_0)$$

for all $v \neq v_0$ in M_K.

Proof. Let c_1 be the number of Theorem 2. Let $c_0 = 1$ if v_0 is archimedean, and let $c_0 = \mathbf{N}\mathfrak{p}_0$, if $\mathfrak{p}_0$ is the prime of v_0. Let $\mathfrak{d}'$ be an M_K-divisor which differs from $\mathfrak{d}$ only at v_0 (i.e. $\mathfrak{d}'(v) = \mathfrak{d}(v)$ for $v \neq v_0$) and such that

$$1/c_1 \leqq \|\mathfrak{d}'\|_K \leqq c_0/c_1 .$$

If v_0 is archimedean, we can adjust the v_0-component as we please, in a continuous fashion. If v_0 is discrete, the value group ranges over powers of $\mathbf{N}\mathfrak{p}_0$, and so we can also find the $\mathfrak{d}'$ subject to our condition. We set $c(v_0) = c_0/c_1$.

By Theorem 2, $\lambda(\mathfrak{d}') > 1$ and hence there exists an element $\alpha \neq 0$ in $L(\mathfrak{d}')$, that is, $\| \alpha \|_v \leqq \| \mathfrak{d}' \|_v$ for all $v \in M_K$. We put $\beta = 1/\alpha$. Then the inequality on the left in the lemma is clearly satisfied. For the inequality on the right, we have

$$\| \beta\mathfrak{d}' \|_v = \frac{\| \beta\mathfrak{d}' \|_K}{\prod\limits_{w \neq v} \| \beta\mathfrak{d}' \|_w} \leqq \| \beta\mathfrak{d}' \|_K = \| \mathfrak{d}' \|_K$$

for all $v \in M_K$. The product is taken over all $w \in M_K$, $w \neq v$. Since $\mathfrak{d}'$ is like $\mathfrak{d}$ except at v_0, we have also proved the inequality on the right.

Returning to the main proof, we note that if $v \in S$, then the value group of v is infinite cyclic, generated by $\mathbf{N}\mathfrak{p}$, and there is only a finite number of primes such that $\mathbf{N}\mathfrak{p} \leqq c(v_0)$. Consequently, there is a finite set of absolute values $S' \supset S$ such that for $v \notin S'$, we have

$$1 = \| \beta\mathfrak{d} \|_v = \| \beta \|_v$$

in the lemma. Furthermore, there is only a finite number of possible sets of values

$$\| \beta\mathfrak{d} \|_v = \| \beta \|_v \qquad v \in S' - S\,.$$

Let $\beta_1, \cdots, \beta_r$ realize this finite number. Given any $\mathfrak{d}$ and any β satisfying the relation in the lemma, there exists some β_j such that

$$\| \beta_j \|_v = \| \beta \|_v$$

for all $v \in S' - S$. This implies that $\beta = \beta_j u$ with some $u \in K_S$. Substituting in the lemma, we get

$$\| u\mathfrak{d} \|_v \leqq c(v_0)/\| \beta_j \|_v$$

for all $v \neq v_0$ in M_K. As the $\beta_1, \cdots, \beta_r$ are fixed, we vary $\mathfrak{d}$ so that $\mathfrak{d}(v) = 1$ for $v \notin S$ and $\mathfrak{d}(v)$ is very large for all $v \in S$. Then $\| u \|_v$ is very small for all $v \in S$ and $v \neq v_0$. By the product formula $\| u \|_{v_0}$ must be large. This achieves the first of our objectives.

(The reader will observe that the crucial point was the finiteness of the $\mathfrak{p}$ with $\mathbf{N}\mathfrak{p}$ bounded. Otherwise, the argument is formal, and applies for instance to function fields in one variable over a finite constant field. We leave this as an exercise.)

As to the second, we have found $x_1, \cdots, x_{s-1} \in K_S$ such that

$$\log x_1 = (\xi_{11}, \cdots, \xi_{1s})$$
$$\cdots$$
$$\log x_{s-1} = (\xi_{s-1,1}, \cdots, \xi_{s-1,s}),$$

and such that the matrix of signs of the ξ_{ij} is as follows:

$$\begin{bmatrix} + & - & - & \cdots & - & - \\ - & + & - & \cdots & - & - \\ & & & \cdots & & \\ - & - & - & \cdots & + & - \end{bmatrix}$$

Let $Y_1, \cdots, Y_s$ be the column vectors. We must show that the first $s-1$ are linearly independent over $\mathbf{R}$. Suppose that

$$a_1 Y_1 + \cdots + a_{s-1} Y_{s-1} = 0$$

not all the coefficients being 0. Say $a_1 > 0$ and $a_1 \geqq a_j$ for any j. Then looking at the sum just in the first row, we get

$$\begin{aligned} 0 &= a_1\xi_{11} + a_2\xi_{12} + \cdots + a_{s-1}\xi_{1,s-1} \\ &\geqq a_1\xi_{11} + a_1\xi_{12} + \cdots + a_1\xi_{1,s-1} \\ &= a_1(\xi_{11} + \xi_{12} + \cdots + \xi_{1,s-1}) \end{aligned}$$

because ξ_{1j} is negative for $j = 2, \cdots, s-1$. By the product formula we must have $\xi_{11} + \xi_{12} + \cdots + \xi_{1,s-1} > 0$, a contradiction.

For the convenience of the reader, we repeat the proof that a discrete subgroup of Euclidean space is a free abelian group. We do this by induction on the dimension of the subgroup, i.e. the maximal number of linearly independent elements over R.

Let Γ be our subgroup and $\xi_1, \cdots, \xi_m$ a maximal set of independent vectors in Γ. Let Γ_0 be the subgroup of Γ contained in the subspace spanned by $\xi_1, \cdots, \xi_{m-1}$. By induction, we may assume that any vector of Γ_0 is a linear integral combination of $\xi_1, \cdots, \xi_{m-1}$.

Consider the subset $\mathfrak{S}$ of all $\xi \in \Gamma$ of the form

$$\xi = a_1\xi_1 + \cdots + a_m\xi_m$$

with real coefficients a_i satisfying

$$0 \leqq a_i < 1 \qquad i = 1, \cdots, m-1$$
$$0 \leqq a_m \leqq 1 .$$

It is a bounded set. Let ξ'_m be a vector of $\mathfrak{S}$ with the smallest possible $a_m \neq 0$, say

$$\xi'_m = b_1\xi_1 + \cdots + b_m\xi_m \, .$$

Starting with any vector ξ of Γ, we can select integral coefficients $c_1, \cdots, c_m$ in such a way that

$$\xi' = \xi - c_m\xi'_m - c_1\xi_1 - \cdots - c_{m-1}\xi_{m-1}$$

lies in $\mathfrak{S}$, and the coefficient of ξ_m is less than b_m and $\geqq 0$. This coefficient must therefore be 0, and ξ' lies in Γ_0. From this our result is clear.

§ 4. *Relative units and ideal classes*

The results of this section will not be used anywhere else in this book, except in Theorem 8 *of Chapter* VII, § 4.

Let K be a field with a proper set of discrete valuations M. We denote by U_M, D_M, C_M the associated group of units, divisors (or ideals, same thing in this case), and divisor classes. We then have by definition, an exact sequence

$$0 \to U_M \to K^* \to D_M \to C_M \to 0 \, .$$

Let M' be a subset of M. Then we have the obvious surjection

$$D_M \to D_{M'} \to 0$$

whose kernel is $D_{M-M'}$. Hence we have the following exact and commutative diagram:

$$\begin{tikzcd}
 & & & 0 \arrow[d] & & \\
 & & & D_{M-M'} \arrow[d] & & \\
0 \arrow[r] & U_M \arrow[r] & K^* \arrow[r] & D_M \arrow[r] \arrow[d] & C_M \arrow[r] & 0 \\
0 \arrow[r] & U_{M'} \arrow[r] & K^* \arrow[r] & D_{M'} \arrow[r] \arrow[d] & C_{M'} \arrow[r] & 0 \\
 & & & 0 & &
\end{tikzcd} \tag{2}$$

From this one gets in the usual manner two exact sequences:

$$0 \to U_M \to U_{M'} \to D_{M-M'} \tag{3}$$

$$D_{M-M'} \to C_M \to C_{M'} \to 0 \tag{4}$$

obtained by trivial chasing around the diagram, which we shall leave to the reader. As an application, we get:

PROPOSITION 5. *Suppose that M' differs from M by a finite number of elements. Then U_M (resp. C_M) is finitely generated if and only if $U_{M'}$ (resp. $C_{M'}$) is finitely generated.*

THEOREM 4. *Let k be a field, V a normal variety scheme over k, and M the set of discrete valuations of $k(V)$ corresponding to the prime divisors of V rational over k. Let U_M be its set of units in $k(V)^*$. Then U_M/k^* is finitely generated.*

Proof. If V is complete, then $U_M = k^*$ and our assertion is trivial. In general, by considering an affine open subset of V whose set of prime rational divisors differs from that of V by a finite set, we may assume that V is affine. Then one may embed V in a projective k-normal variety $\bar{V}$ whose set of prime divisors over k again differs from that of V by a finite set. We can now use the exact sequence (3), modified as follows:

$$0 \to U_M/k^* \to U_{M'}/k^* \to D_{M-M'}$$

to conclude the proof.

COROLLARY. *Let k be a field, K a finitely generated regular extension of k, and S a subalgebra of K finitely generated as an algebra over k. Let S^* be the invertible elements of S. Then S^*/k^* is a finitely generated group.*

Proof. Our algebra determines a variety, and by going to an affine open subset of this variety corresponding to a larger ring S_1, we may assume that this variety is non-singular in codimension 1. (Indeed, S^*/k^* is canonically contained in S_1^*/k^*.) The invertible elements of S_1 coincide with the units for its set of prime divisors, and we can apply the theorem.

Let R be a Noetherian ring, integrally closed, let V be a variety scheme over R and assume that V is normal. Using the notation of §1, let $M'_R = M'_{R,V}$ be the subset of $M_{R,V}$ consisting of those elements of M_R having a unique extension to M_V, and let $M'_V = M'_R \cup M_{k,V}$ where k is the quotient field of R. Then we have an exact and commutative diagram:

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
 & U_{M'_R} \arrow[d] & U_{M'_V} \arrow[d] & U_{M_{k,V}}/k^* \arrow[d] & \\
0 \arrow[r] & k \arrow[r] \arrow[d] & K^* \arrow[r] \arrow[d] & K^*/k^* \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & D_{M'_R} \arrow[r] \arrow[d] & D_{M'_V} \arrow[r] \arrow[d] & D_{M_{k,V}} \arrow[r] \arrow[d] & 0 \\
 & C_{M'_R} \arrow[d] & C_{M'_V} \arrow[d] & C_{M_{k,V}} \arrow[d] & \\
 & 0 & 0 & 0 &
\end{tikzcd} \tag{5}$$

and consequently an exact sequence

$$U_{M'_R} \to U_{M'_V} \to U_{M_{k,V}}/k^* \to C_{M'_R} \to C_{M'_V} \to C_{M_{k,V}} \tag{6}$$

where the middle map is the usual coboundary operator (cf. [3]).

If we now apply Theorem 4 to the first part of this sequence, together with the finite generation of the unit group in number fields, we get:

THEOREM 5. *Let V be a normal variety scheme over the ring of integers I of an algebraic number field k. Then the group of units U_{M_V} of its field of rational functions is finitely generated.*

COROLLARY. *Let R be a finitely generated ring over* $\mathbf{Z}$. *Then the invertible elements R^* of R form a finitely generated group.*

Proof. Using Proposition 2 of §1, we may embed R in an integrally closed ring of finite type over $\mathbf{Z}$, and it clearly suffices to prove our assertion in this case. Our ring then determines a variety scheme as in the theorem, and its invertible elements are equal to the M_V-units, so that the theorem applies.

It will be a comparatively deep theorem that the group $C_{M_{k,V}}$ is finitely generated whenever k is a field of finite type over the prime field. Once this will be proved, however, (Chapter V) we shall see, from the end of the exact sequence (6) together with the finiteness of class number theorem in number fields (Theorem 1 of §2), that we can derive a similar statement for the divisor classes:

THEOREM 6. *Let V be a normal projective variety scheme over the ring of integers of a number field k. Let M_V be the set of discrete* *valuations corresponding to its prime divisors, as in Example* 3 *of* §1. *Then the group C_{M_V} of divisor classes is finitely generated.*

We apply sequence (6) using the set M_V' which differs from M_V by a finite subset. This of course makes no difference to the finite generation, in view of sequence (4).

Historical Note:

It was Artin-Whaples [2] who first showed how the product formula can be taken as the basic axiom for algebraic number theory. For us here, it is irrelevant to know that if one of the primes is archimedean or discrete and its residue class field is finite over the constant field, then this characterizes the field as being either a number field or a function field in one variable.

Divisors, including possibly archimedean primes, were defined by Weil [40], with the idea of making his decomposition theorem valid for them also. We shall reproduce this theorem in Chapter VII.

The theorem giving the approximate number of algebraic numbers (Theorem 2), as well as the proof of the unit theorem, has been taken from Artin-Whaples. In this respect, one must emphasize that the finiteness of the ideal class group, and the finite generation of the unit group are utter trivialities. In the case of the units, this is *not* the content of the unit theorem. Its content lies first in giving the precise rank, but even more in asserting that the log vectors, as in Theorem 3, are linearly independent over **R** which is saying much more. In some applications, it is necessary to have the unit theorem as we have given it. However, in §4, or for applications to Chapter VI, where we are interested only in qualitative results, the finite generation would suffice. For the ideal class group, there is an analogue to the quantitative result which we have given here for the unit group, i.e. the more precise determination of the Minkowski constant of Theorem 1. We shall not go into this here.

One should note that Theorem 2 can be improved to give an asymptotic estimate for the number of elements in K in $L(\mathfrak{d})$, namely, we have

$$\lambda(\mathfrak{d}) = c_K \,||\, \mathfrak{d} \,||_K + o(||\, \mathfrak{d} \,||_K)$$

for $||\, \mathfrak{d} \,||_K \to \infty$. This is more or less known, but is not explicitly in the literature. One reduces the proof to counting lattice points in parallelotopes in Euclidean N-space (where $N = [K : \mathbf{Q}]$) by means of the finiteness of the class number and Lemma 2.

One can then use a simple direct argument. We shall not go further into this here.

The contents of § 4 are taken from Roquette [30].

Finally we observe that there are product formulas arising from situations more general than that of Example 2. Indeed, let us take a complete, non-singular surface V and a curve C on V which has the following property: For every positive 1-cycle X on V, the intersection number $I(X.C)$ of the numerical equivalence classes of X and C is $\geqq 0$. For every prime rational cycle $\mathfrak{p}$ on V we define the multiplicity $d(\mathfrak{p}) = I(\mathfrak{p}.C)$. If φ is a function on V, then we define

$$|\varphi| = e^{-d(\mathfrak{p}) \operatorname{ord}_{\mathfrak{p}}(\varphi)}.$$

The set of such absolute values does satisfy the product formula. If the curve is a generic hyperplane section, then we have our standard example. However, as Mumford pointed out to me, there exist examples of such curves which do not lie in a pencil, and thus give rise to other types of product formulas, to which Roth's theorem would also apply.

CHAPTER III

Heights

The possibility of defining the height of a point on a variety lies at the base of all possibilities of counting such points. In this book, this allows us to get qualitative results, to the effect that certain sets of points are finite, or, if they form a group, a finitely generated one.

One can give the definition of the height just after the axiom that we have a field with a proper set of absolute values satisfying the product formula. We can also prove important geometric properties of the height (how it varies under certain mappings) using only this axiom. The formalism we shall develop will then be used in the infinite descent of Chapter V, and in the counting of Siegel's theorem in Chapter VII.

In number fields, the set of points in a given projective space of bounded degree and bounded height is finite. We shall give in §3 the analogue in function fields, where we prove that certain varieties lie in a finite number of algebraic families.

§1. *Definitions*

Let **F** be a field with a proper set of absolute values $M_{\mathbf{F}}$, satisfying the product formula. Then for each finite extension K of **F**, the set of absolute values on K extending those of $M_{\mathbf{F}}$ is a proper set M_K, satisfying the product formula with multiplicities $N_v = [K_v : \mathbf{F}_v]$ for $v \in M_K$. As before, we set

$$\| x \|_v = | x |_v^{N_v} \qquad x \in K .$$

We now consider projective n-space $\mathbf{P}^n$, and let $P \in \mathbf{P}^n_K$ be a point in projective space, rational over K. Let $(x_0, \cdots, x_n)$ be a set of coordinates for P, with each $x_i \in K$. Then we define the *height of P relative to* M_K by the formula

$$H_{M_K}(P) = \prod_{v \in M_K} \sup_i \| x_i \|_v .$$

In view of the product formula, one sees immediately that our product is independent of the choice of coordinates chosen: If

$y \in K^*$ then

$$\sup_i \| y x_i \|_v = \| y \|_v \sup_i \| x_i \|_v$$

and taking the product, the term involving y will be equal to 1.

In particular, we define the height of an element $x \in K$ to be the height of the point $(1, x)$ in $\mathbf{P}^1$, so that we have

$$H_{M_K}(x) = \prod_{v \in M_K} \sup (1, \| x \|_v) ,$$

and we see that if $x \neq 0$, then $H_{M_K}(x) = H_{M_K}(1/x)$.

For convenience, if the set $M_{\mathbf{F}}$ is fixed throughout, we write simply H_K or H, if the reference to K is also fixed throughout a discussion.

Let us suppose that our set $M_{\mathbf{F}}$ is fixed. Let $E \supset K$ be two finite extensions of $\mathbf{F}$, and let $(x_0, \cdots, x_n)$ be coordinates of a point P in $\mathbf{P}^n$, rational over K. Then one has

$$H_E(P) = H_K(P)^{[K:E]} .$$

Indeed, letting $N_w = [E_w : \mathbf{F}_w]$ for $w \in M_E$, we get

$$H_E(P) = \prod_{w \in M_E} \sup_i | x_i |_w^{N_w} = \prod_{v \in M_K} \prod_{w | v} \sup_i | x_i |_w^{N_w} .$$

On the other hand,

$$[E_w : \mathbf{F}_w] = [E_w : K_w][K_w : \mathbf{F}_w]$$

For all absolute values w inducing the same absolute value v on K, the term $[K_w : \mathbf{F}_w]$ is constant, and we have

$$\sum_{w | v} [E_w : K_w] = [E : K] .$$

From this our assertion is obvious.

The formula we have just proved shows how the height changes as we extend the field. It is therefore natural to define an *absolute height* of a point P, rational over the algebraic closure of $\mathbf{F}$, by

$$h(P) = h_M(P) = H_K(P)^{1/[K:\mathbf{F}]}$$

for any finite extension K of $\mathbf{F}$ over which P is rational. The argument we have just given shows that $h(P)$ is independent of the choice of this field.

Let σ be an isomorphism of K over $\mathbf{F}$ (i.e. leaving $\mathbf{F}$ fixed). Let P be a point as above, with coordinates $(x_0, \cdots, x_n)$ rational over K. Then we can define the point P^σ, rational over K^σ, and having coordinates $(x_0^\sigma, \cdots, x_n^\sigma)$. By transport of structure, we get immediately

$$H_K(P) = H_{K^\sigma}(P^\sigma)$$

whence in particular,

$$h(P) = h(P^\sigma) .$$

Let us now consider an example. We take $\mathbf{F} = \mathbf{Q}$, the rational numbers. We can write any non-zero rational number in the form a/b where a, b are ralatively prime integers. If p is a prime number then $|a|_p$ and $|b|_p$ are $\leqq 1$, and one of them $=1$. Taking the canonical set of absolute values on $\mathbf{Q}$ to define the height, we see that

$$H_{\mathbf{Q}}(a/b) = \sup(|a|, |b|) .$$

More generally, any point in projective n-space rational over $\mathbf{Q}$ has a set of coordinates $(x_0, \cdots, x_n)$ which are relatively prime integers, and we then see that

$$H_{\mathbf{Q}}(P) = \sup_i |x_i|$$

the absolute value being the ordinary one. In particular, the set of points in $\mathbf{P}^n$ rational over $\mathbf{Q}$, and of height $\leqq$ a fixed number is finite. We shall generalize this later to a number field. *We agree from now on that the heights in number fields will always be those relative to the canonical set of absolute values on* $\mathbf{Q}$.

§ 2. *Gauss' lemma*

Let K be a field with a non-trivial absolute value v. If $f(X) = \sum a_\nu \mathbf{M}_\nu(X)$ is a polynomial in several variables with coefficients $a_\nu \in K$, and

$$\mathbf{M}_\nu(X) = \mathbf{M}_\nu(X_1, \cdots, X_n) = X_1^{\nu_1} \cdots X_n^{\nu_n}$$

a monomial, then we define

$$|f|_v = |f| = \sup_\nu |a_\nu| .$$

Gauss' Lemma for valuations then asserts that this is a valuation if v is a valuation:

Proposition 1. *Let K be a field with a non-trivial valuation. If $f, g \in K[X_1, \cdots, X_n]$ then*

$$|fg| = |f|\,|g|\,.$$

Proof: Let us first assume that f, g are polynomials in one variable, so we can write

$$f(X) = a_d X^d + \cdots + a_0$$
$$g(X) = b_e X^e + \cdots + b_0\,.$$

Say a_r is a coefficient of f such that $|a_r| \geqq a_i$ for all i, and is the one which lies furthest to the left that has this property. Similarly for b_s with respect to the b_j. Then

$$(1/a_r)f = (a_d/a_r)X^d + \cdots + X^r + \cdots + a_0/a_r$$
$$(1/b_s)g = (b_e/b_s)X^e + \cdots + X^s + \cdots + b_0/b_s\,,$$

and the coefficients of these polynomials each have a value $\leqq 1$. Those to the left of X^r (resp. X^s) have a value <1. If we take the product

$$(1/a_r b_s)fg$$

then we see that $X^r X^s$ has a coefficient of type $1 + c$ where $|c| < 1$. The coefficient of every power X^i with $i > r + s$ has a value <1, while if $i < r + s$, then its value is $\leqq 1$. From this one sees that

$$|(1/a_r b_s)fg| = 1$$

and hence that $|fg| = |a_r b_s|$, which proves our proposition for polynomials in one variable.

Now let f be a polynomial in n variables $X_1, \cdots, X_n$, of degree $<d$. Then the polynomial in one variable

$$(S_d f)(Y) = f(Y, Y^d, \cdots, Y^{d^{n-1}})$$

has the same set of non-zero coefficients as f. Thus, if f and g are two polynomials in n variables, such that the sum of their degrees is $<d$, then

$$S_d(fg) = S_g(f)S_d(g)$$

has the same non-zero coefficients as fg. From this our reduction of the n-variable case to the 1-variable case is clear.

For an archimedean absolute value, we do not get the full multiplicativity, but we shall see that we get a modified version of it. Let us first consider the case of one variable.

LEMMA 1. *Let K be a field of characteristic 0, with an absolute value which coincides with the ordinary one on $\mathbf{Q}$. Let $f(X) \in K[X]$ be a polynomial of degree d, and let*

$$f(X) = \prod_{i=1}^{d} (X - \alpha_i)$$

be a factorization in $\bar{K}$. We assume that our absolute value is extended to $\bar{K}$. Then

$$\frac{1}{2^d} \prod_{i=1}^{d} \sup(1, |\alpha_i|) \leqq |f| \leqq 2^d \prod_{i=1}^{d} \sup(1, |\alpha_i|) .$$

Proof: The right inequality is trivially proved by induction, estimating the coefficients in a product of a polynomial $g(X)$ by $(X - \alpha)$. We prove the other by induction on the number of indices i such $|\alpha_i| > 2$. If $|\alpha_i| \leqq 2$ for all i, our assertion is obvious. Suppose now that $f(X) = g(X)(X - \alpha)$ with $|\alpha| > 2$ and suppose that our assertion is true for $g(X) = X^d + b_{d-1}X^{d-1} + \cdots + b_0$. We have

$$f(X) = X^{d+1} + (b_{d-1} - \alpha)X^d + (b_{d-2} - \alpha b_{d-1})X^{d-1} + \cdots + (-\alpha)b_0 .$$

We can assume $|g| = |b_i| \geqq |b_{i-1}|$ for some i, $0 \leqq i \leqq d$ (with the convention $b_d = 1, b_{-1} = 0$). Then

$$\begin{aligned} |f| \geqq |\alpha b_i - b_{i-1}| &\geqq |\alpha||b_i| - |b_{i-1}| \\ &\geqq |\alpha||b_i| - |b_i| = (|\alpha| - 1)|b_i| \\ &\geqq |\alpha|/2 \cdot |b_i| = |\alpha|/2 \cdot |g| \end{aligned}$$

and our lemma is now obvious, since $|\alpha| > 2$.

PROPOSITION 2. *Let K be a field of characteristic 0, with an absolute value which coincides with the ordinary one on $\mathbf{Q}$. Let d be an integer $\geqq 0$. If f and g are two polynomials in $K[X_1, \cdots, X_n]$ such that deg f + deg $g < d$, then*

$$\frac{1}{4^{dn}} |fg| \leqq |f||g| \leqq 4^{dn} |fg| .$$

Proof: Using the substitution S_d as before, we are reduced to the case of one variable with $n = 1$. Say

$$f(X) = a \prod (X - \alpha_i)$$
$$g(X) = b \prod (X - \beta_j) .$$

Without loss of generality, we may assume $a = b = 1$, and that we have extended our absolute value to $\bar{K}$. Using the lemma, we get

$$\begin{aligned} |f||g| &\leqq 2^d \prod \sup(1, |\alpha_i|) \prod \sup(1, |\beta_j|) \\ &\leqq 2^d 2^d |fg| . \end{aligned}$$

The other inequality is proved in the same way.

Let $\mathbf{F}$ again be a field with a proper set of absolute values $M_{\mathbf{F}}$ satisfying the product formula. We consider a polynomia

$$f(X) = \sum a_j \mathbf{M}_j(X)$$

in several variables, with coefficients in $\bar{\mathbf{F}}$, and thus in a finite extension K of $\mathbf{F}$. We define its *absolute height* $h(f)$ to be the height $h(P)$ of the point P having the a_j (in any order) as coordinates, and we define its *relative height* $H_K(f)$ in a similar way.

From Proposition 2 we can now deduce analogous results for heights.

PROPOSITION 3. *Let* $\mathbf{F}$ *be a field with a proper set of absolute values* $M_{\mathbf{F}}$ *satisfying the product formula, and let* s *be the number of archimedean absolute values in* $M_{\mathbf{F}}$. *Let* f *and* g *be two polynomials in* n *variables with coefficients in* $\mathbf{F}$, *with* $\deg f + \deg g < d$. *Then*

$$\frac{1}{4^{d^n s}} h(fg) \leqq h(f)h(g) \leqq 4^{d^n s} h(fg) .$$

Proof: Let $c_1 = 1/4^{d^n}$ and $c_2 = 4^{d^n}$. Let K be a finite extension of F in which f and g have their coefficients. For $v \in M_K$, put $||f||_v = |f|_v^{N_v}$. We have

$$H_K(fg) = \prod_{v \in M_K} ||fg||_v \geqq \prod_{v \in M_K} ||f||_v \, ||g||_v \; c_v$$

where $c_v = 1$ if v is non-archimedean, and otherwise, $c_v = c_1^{N_v}$. Let S_K be the set of archimedean absolute values in M_K. Then

$$\sum_{v \in S_K} N_v = [K : \mathbf{F}]\, s$$

whence

$$\prod_{v \in M_K} c_v = c_1^{[K:\mathbf{F}]s}$$

and the inequality on the left follows immediately. The one on the right follows in a similar way.

Let $f(X)$ be a as above, with coefficients in a finite extension K of $\mathbf{F}$. If σ is an isomorphism of K over $\mathbf{F}$ then we get the polynomial

$$f^{\sigma}(X) = \sum a_i^{\sigma} \mathbf{M}_j(X) ,$$

and thus, as for points, we have

$$h(f) = h(f^{\sigma}) .$$

Let α be algebraic over $\mathbf{F}$, and let $f(X)$ be its irreducible polynomial over $\mathbf{F}$. Then

$$f(X) = \prod_{j=1}^{d} (X - \alpha_j)$$

where d is the degree of α over $\mathbf{F}$ (i.e. $[\mathbf{F}(\alpha) : \mathbf{F}]$) and the α_j are the conjugates of α, each repeated with a suitable multiplicity in characteristic p. In view of the above results we get:

Proposition 4. *Let $\mathbf{F}$ be a field with a proper set of absolute values satisfying the product formula. Let d be an integer >0. Then there exist two numbers c_1, $c_2 > 0$ depending on d, such that if α is algebraic over $\mathbf{F}$ of degree d, and $f(X)$ is its irreducible polynomial over $\mathbf{F}$, then*

$$c_1 h(\alpha)^d \leqq h(f) \leqq c_2 h(\alpha)^d .$$

Let us consider the special case where $\mathbf{F} = \mathbf{Q}$ is the rational numbers. Then $h(f)$ is the maximum of the absolute values of the coefficients of the irreducible polynomial of α over $\mathbf{Z}$. More generally, consider a point $(\alpha_0, \cdots, \alpha_n)$ in projective space, rational over a finite extension K of $\mathbf{Q}$, of degree d. Consider the polynomial

$$f(X) = \alpha_0 X_0 + \cdots + \alpha_n X_n .$$

Then $h(f)$ is the height of our point. Let

$$g(X) = \prod_{\sigma} f^{\sigma}$$

the product being taken over all distinct isomorphisms σ of K over $\mathbf{Q}$. Then g has coefficients in $\mathbf{Q}$, and $h(g)$, $h(f)^d$ have the same order of magnitude. We have already seen in § 1 that the number of points with height less than a fixed number, in a projective space, and rational over $\mathbf{Q}$ is bounded. Consequently we now get:

Theorem 1. *Let d_0, h_0 be two fixed numbers >0. Then the set of points P in $\mathbf{P}^n$ algebraic over $\mathbf{Q}$, and such that*

$$[\mathbf{Q}(P) : \mathbf{Q}] < d_0 \text{ and } h(P) < h_0$$

is finite.

We note the strong uniformity in the statement of the above proposition. Not only is the set of points in $\mathbf{P}^n$ rational over a *given* number field and of bounded height finite, but so is the set of points of bounded degree and bounded height. Actully, in the final applications, this stronger uniformity will not be used, because we shall consider only points of varieties which are rational over a fixed number field (for instance in the Mordell-Weil theorem, or the Siegel theorem).

§ 3. *Heights in function fields*

We shall consider the height function from a geometric point of view.

Let W be a projective variety in $\mathbf{P}^r$, non-singular in codimension 1, and defined over a field k. Let c be a number, $0 < c < 1$. Let $M_K = M_{k(W)}$ be the set of discrete absolute values of the functon field $k(W)$ obtained from the prime rational divisors of W over k. We then have by definition

$$| x |_{\mathfrak{p}} = c^{(\operatorname{ord}_{\mathfrak{p}} x) \deg(\mathfrak{p})}$$

for each such prime divisor $\mathfrak{p}$, and our set M_K satisfies the product formula. If P is a point in $\mathbf{P}^n$ rational over $k(W)$ with coordinates $(y_0, \cdots, y_n)$ in $k(W)$, then

$$H_{k(W)}(P) = H_W(P) = \prod_{\mathfrak{p}} \sup_i | y_i |_{\mathfrak{p}} .$$

Proposition 5. *Let* $d = \deg \sup_i \, (y_i)_\infty$ *be the projective degree in* $\mathbf{P}^r$ *of the sup of the polar divisors of the* y_j. *Then*

$$H_W(P) = (1/c)^d .$$

Proof: Without loss of generality, we may assume that one of the y_i is equal to 1. In our product defining the height, we will have a contribution $\neq 1$ only if $\mathfrak{p}$ is a pole of one of the y_i. For such a pole $\mathfrak{p}$ we have

$$\begin{aligned} \sup_i c^{(\operatorname{ord}_{\mathfrak{p}} y_i)\deg(\mathfrak{p})} &= \sup_i (1/c)^{(-\operatorname{ord}_{\mathfrak{p}} y_i) \deg(\mathfrak{p})} \\ &= (1/c)^{\sup_i(-\operatorname{ord}_{\mathfrak{p}} y_i)\deg(\mathfrak{p})} . \end{aligned}$$

Our assertion is now obvious.

We shall call $d = \deg \sup_i (y_i)_\infty$ the *logarithmic height*, (or simply height if we work with it throughout a discussion), and we shall denote it by $\eta_W(P)$. It can also be interpreted in the following manner.

Proposition 6. *Let W be a projective variety in $\mathbf{P}^m$, non-singular in codimension 1, defined over the field k, and let P be a point in projective space $\mathbf{P}^n$, rational over $k(W)$. Let $f: W \to \mathbf{P}^n$ be the rational map defined over k, determined by P. Then*

$$\eta_W(P) = \deg f^{-1}(L)$$

for any hyperplane L of $\mathbf{P}^n$, such that $f^{-1}(L)$ is defined, the degree being that in the given projective embedding of W in $\mathbf{P}^m$.

Proof: Without loss of generality, we may assume none of the coordinates $(y_0, \cdots, y_n)$ of P is 0. Let $(Y_0, \cdots, Y_n)$ be the variables of $\mathbf{P}^n$. Each $Y_i/L(Y)$ determines a function φ_i on $\mathbf{P}^n$, i.e. a rational map $\varphi_i: \mathbf{P}^n \to \mathbf{P}^1$. We have two successive maps:

$$W \xrightarrow{f} \mathbf{P}^n \xrightarrow{\varphi_i} \mathbf{P}^1$$

and if $L(y) \neq 0$ (i.e., $f^{-1}(L)$ is defined) then $y_i/L(y)$ determines the composition function $\varphi_i \circ f = \psi_i$ of W into $\mathbf{P}^1$. By AV, Corollary 1 of Proposition 2, Appendix, § 1, we get

$$(\psi_i) = f^{-1}(L_i) - f^{-1}(L) ,$$

where L_i is the hyperplane $Y_i = 0$ of $\mathbf{P}^n$. Since the L_i are without common point, the divisors $f^{-1}(L_i)$ are without common component, and hence

$$\sup_i(\psi_i)_\infty = f^{-1}(L) .$$

This proves our proposition.

We observe that our logarithmic height, and thus the height, are geometric, i.e. essentially do not depend on the field of definition of W. In particular, our point P being also rational over $k(W)$, we have

$$H_{\bar{k}(W)}(P) = H_{k(W)}(P) .$$

Let W be as before, and let P be a point in $\mathbf{P}^n$ rational over $k(W)$. Then P defines a generically surjective rational map

$$g: W \to T$$

of W, such that if $j: T \to \mathbf{P}^n$ is the inclusion, then $j \circ g = f$. If x is a generic point of W over k, we have $f(x) = g(x)$. This is a generic point of T over k.

Since T is contained in $\mathbf{P}^n$, it has a degree in $\mathbf{P}^n$ which we shall also write deg T or $\deg_{\mathbf{P}^n} T$ if there is any chance of confusion with respect to the embedding. We wish to give a rough comparison of $\eta_W(P)$ and deg T. To do this, we need some lemmas.

We start therefore with a generically surjective rational map $g : W \to T$ defined over k. Let dim $T = s$, and let $L_{u_1}, \cdots, L_{u_s}$ be generic independent hyperplanes of $\mathbf{P}^n$ over k. This means that $u_1, \cdots, u_s$ are independent generic over k (cf. IAG, Chapter VIII, § 6). We also write L_i instead of L_{u_i}. Then $T \cdot L_1 \cdots L_s$ is a cycle on $\mathbf{P}^n$ consisting of deg T points, which are all generic points of T over k, by IAG, Proposition 10 of Chapter VIII, § 6. Furthermore, each one of these points is separable over $k(u_1, \cdots, u_s)$, say by the criterion for multiplicity 1.

If y is a generic point of T over k and $y = f(x) = g(x)$, then $g^{-1}(y)$ is the prime rational cycle on W equal to the locus of x over $k(y)$, taken with multiplicity 1 by F^1-VII_6 Th. 12, i.e. it is the "generic fiber". (It is not necessarily a variety, i.e. absolutely irreducible.)

LEMMA 2. *Let W be a projective variety in $\mathbf{P}^m$. Let $g : W \to T^s$ be a generically surjective rational map defined over k. Let y be a generic point of T over k, and let $U = g^{-1}(y)$. Let $j : T \to \mathbf{P}^n$ be an inclusion, and let $L_1 \cdots L_s$ be generic independent hyperplanes of $\mathbf{P}^n$ over k. Then*

$$(\deg T)(\deg U) = \deg g^{-1}(T \cdot L_1 \cdots L_s)\,.$$

Proof: Let $T \cdot L_1 \cdots L_s = \sum (y_i)$. Then deg T is the number of these points (y_i), and the cycle on the right is precisely $\sum U_i = \sum g^{-1}(y_i)$. The U_i are all conjugate over k (i.e. each one can be transformed into the other by an automorphism of the universal domain) and hence have the same degree in $\mathbf{P}^m$, namely the degree of U. Our assertion is therefore clear.

LEMMA 3. *Let the notation be as in the previous lemma, and $f = j \circ g$. Suppose that y is one of the points of $T \cdot L_1 \cdots L_s$. Then each component of $U = g^{-1}(y)$ is a proper component of*

$$f^{-1}(L_1) \cap \cdots \cap f^{-1}(L_s)$$

and we have

$$g^{-1}(T \cdot L_1 \cdots L_s) \leqq f^{-1}(L_1) \cdots f^{-1}(L_s)\,.$$

Proof: Let U_0 be a component of U and let U_0' be a component of $f^{-1}(L_1) \cap \cdots \cap f^{-1}(L_s)$ containing U_0 (which is itself contained in their intersection). Observe that f and g are both defined along U_0 (since a generic point for U_0 is also one for U over $k(y)$ and is a generic point for W over k). Hence f and g are defined along U_0'. Let x' be a generic point of U_0' over the algebraic closure of $k(u_1, \cdots, u_s)$ (notation as in the beginning of the discussion). Then $f(x')$ lies in $T \cap L_1 \cap \cdots \cap L_s$ and hence is one of the points y of $T \cdot L_1 \cdots L_s$. Since $f(x') = g(x')$, it follows that x' lies in supp (U), and thus is one of the components of U. Hence U_0' itself is contained in this component, which must therefore be U_0, and we get $U_0' = U_0$.

To prove the second assertion, note that

$$f^{-1}(L_1) \cdots f^{-1}(L_s)$$

is rational over $k(u_1, \cdots, u_s)$ and as we mentioned before, each point y in $T \cdot L_1 \cdots L_s$ is separable over this field. Since $g^{-1}(y)$ is prime rational over $k(u_1, \cdots, u_s)(y)$ (which is the composite of $k(y)$ and a purely transcendental extension of k, free from $k(y)$ over k), it follows that

$$g^{-1}(y) \leqq f^{-1}(L_1) \cdots f^{-1}(L_s) .$$

If y_1, y_2 are two distinct points y as above, then $g^{-1}(y_1)$ has no component in common with $g^{-1}(y_2)$. Hence we get

$$g^{-1}(T \cdot L_1 \cdots L_s) \leqq f^{-1}(L_1) \cdots f^{-1}(L_s)$$

as desired.

LEMMA 4. *Let W be a projective variety in a projective space $\mathbf{P}^m$. Let $X_1, \cdots, X_s$ be positive divisors on W. Then*

$$\deg(X_1 \cdots X_s) \leqq (\deg X_1) \cdots (\deg X_s) .$$

Proof: Let W and the X_i be rational over an algebraically closed field k. We take a generic linear direction over k as in F^2-IX_3 such that the join of this linear variety with a divisor on V is a divisor on $\mathbf{P}^m$. We call this join also the projecting cone over the divisor, as usual. We recall the following facts. If A is a subvariety of W defined over k, and $\tilde{A}$ the projecting cone over A, then deg $A = \deg \tilde{A}$. If A and B are subvarieties of W defined over k, and $A \neq B$, then $\tilde{A} \neq \tilde{B}$ by F^2-IX_3, Proposition 6, Corollary 3. If A is a proper components of $X_1 \cdots X_s$ of multiplicity m, then $\tilde{A}$ is also a proper component of $\tilde{X}_1 \cdots \tilde{X}_s$ of multiplicity m, by the corollary to Proposition 7 of F^2-IX_3, combined with F^2-$VIII_4$, Theorems 10 and 11. The intersection of the X_i is of course taken on $\mathbf{P}^m$. From these facts, it follows that

$$\deg(X_1 \cdots X_s) \leqq \deg(\tilde{X}_1 \cdots \tilde{X}_s)$$

and this reduces the proof to the case where W is equal to all of projective space.

We then proceed by induction on s. By definition, we have

$$(X_1 \cdots X_s) = (X_1 \cdots X_{s-1}) \cdot X_s \, .$$

Thus it will suffice to prove our assertion for two factors, say X and Y, one of which is a divisor in $\mathbf{P}^m$. If every component of the intersection $X \cap Y$ is proper, we can apply Bezout's theorem. Otherwise, we proceed by induction on the sum of the multiplicities of the components of X and Y. If both X and Y are varieties. with X of codimension 1 in $\mathbf{P}^m$, then either $X \supset Y$ in which case $X \cdot Y = 0$, or $X \not\supset Y$ and Bezout's theorem applies. In any case, our assertion is obvious. Say now $X = X_1 + A$ where A is a variety, and $X_1 > 0$. Then

$$X \cdot Y = (X_1 + A) \cdot Y = X_1 \cdot Y + A \cdot Y \, .$$

By induction, we get

$$\deg(X_1 \cdot Y) + \deg(A \cdot Y) \leqq (\deg X_1)(\deg Y) + (\deg A)(\deg Y)$$

which proves the lemma.

THEOREM 2. *Let W^r be a projective variety, non-singular in codimension* 1 *and defined over a field k. Let P be a point in $\mathbf{P}^n$ rational over $k(W)$, and let T be the locus of P over k, so that we have a generically surjective rational map $g : W \to T$. Then*

$$\deg_{\mathbf{P}^n} T \leqq \eta_W(P)^r \, .$$

Proof: This follows immediately from the preceding lemmas, using $X_i = f^{-1}(L_i)$ and applying Proposition 6.

In particular, we see that if P ranges over a set of points of bounded heights, then T has bounded degree in $\mathbf{P}^n$.

§ 4. ***Heights on abelian groups***

This section is entirely different from the preceding three. In some applications, e.g. to abelian varieties, the height function becomes related to the group law in certain ways, and it is convenient to make an abstraction of this situation. Rather than deal with the height function itself, we prefer to axiomatize the situation for its logarithm. The results proved below are purely combinatorial.

Let Γ be an abelian group. Let d be a real number >0. Let $\eta : \Gamma \to \mathbf{R}$ be a real-valued function on Γ. We shall say that η is of *quasi-degree* d if $\eta(P) \geqq 0$ for all $P \in \Gamma$ and if it satisfies the following condition:

Given an integer $m > 0$, *and* ε *real* > 0, *and an element* $a \in \Gamma$, *there exists a number* $c > 0$ *such that for all* $P \in \Gamma$ *we have*

$$-c + (m^d - \varepsilon)\eta(P) \leqq \eta(mP + a) \leqq c + (m^d + \varepsilon)\eta(P) .$$

We shall say that η is *non-degenerate* if given $\beta > 0$ there exists only a finite number of elements $P \in \Gamma$ such that $\eta(P) \leqq \beta$.

The arguments used to prove the next proposition are known as the *infinite descent.*

PROPOSITION 7. *Let* Γ *be an abelian group and* η *a real valued function on* Γ *of quasi-degree* $d > 0$. *Let* m *be an integer* >1 *such that* $\Gamma/m\Gamma$ *is finite, and let* $a_1, \cdots, a_s$ *be representatives in* Γ *of* $\Gamma/m\Gamma$. *There exists a number* c_1 *and a subset* $\mathfrak{S}$ *of* Γ *such that*:

(i) $\eta(P) \leqq c_1$ *for all* $P \in \mathfrak{S}$.

(ii) *For any* $P_0 \in \Gamma$, *there exist integers* $n_0, n_1, \cdots, n_s$ *and a point* P *in* $\mathfrak{S}$ *such that*

$$P_0 = n_0 P + n_1 a_1 + \cdots + n_s a_s .$$

Proof: In the definition of quasi-degree, the number c depends on m and a, so let us write $c(m, a)$ for it. In the present application, we let $c = \sup_i c(m, a_i)$.

Consider a sequence $(P_0, P_1, \cdots)$ of points of Γ constructed by starting with our point P_0, and such that

$$mP_{\nu+1} = P_\nu - a_{i_\nu} .$$

By definition, there is a number $\lambda > 1$ (in fact, $\lambda = m^d - \varepsilon$) such that

$$\eta(P_{\nu+1}) \leqq \eta(P_\nu)/\lambda + c/\lambda .$$

Proceeding inductively, we get

$$\eta(P_\nu) \leqq 1/\lambda^\nu \cdot \eta(P_0) + c(1 + 1/\lambda + 1/\lambda^2 + \cdots)$$

and from this our assertion is obvious.

COROLLARY. *Let* Γ *be an abelian group and* η *a real-valued function on* Γ *of quasi-degree* $d > 0$, *and non-degenerate. Let* m *be an integer*

>1 *such that* $\Gamma/m\Gamma$ *is finite. Then* Γ *is finitely generated.*

Proof: Immediate consequence of the definition of non-degeneracy. The procedure in the proof of Proposition 7 can be reversed and leads to an infinite ascent, which we use to estimate the number of elements $P \in \Gamma$ such that $\eta(P) \leqq \beta$ as β becomes large.

THEOREM 3. *Let* Γ *be a finitely generated abelian group, of rank* r *(maximum number of linearly independent elements over* $\mathbf{Z}$*). Let* η *be a real-valued function on* Γ*, which is of quasi-degree* $d > 1$ *and non-degenerate. Given* $\varepsilon > 0$ *there exist two numbers* $c_1(\varepsilon), c_2(\varepsilon) > 0$ *and a number* δ *with* $0 < \delta < \varepsilon$ *such that for all* $\beta > \beta_0(\varepsilon)$*, the number* $N(\beta)$ *of elements of* Γ *such that* $\eta(P) \leqq \beta$ *satisfies the inequalities*

$$c_1(\varepsilon)\beta^{r/d-\delta} \leqq N(\beta) \leqq c_2(\varepsilon)\beta^{r/d+\delta} .$$

Proof: We can write Γ as a direct sum of a free abelian group and a finite torsion group. From the definitions, and the approximate nature of our estimates, we may assume without loss of generality that Γ is free over $\mathbf{Z}$.

Let m be an integer $\geqq 2$ and $x_1, \cdots x_r$ a basis for Γ over $\mathbf{Z}$. Let $\{a_j\}$ range over those elements of Γ which can be written in the form

$$n_1x_1 + \cdots + n_rx_r$$

with $-(m-1) \leqq n_i \leqq (m-1)$. By σ_ν we denote a sequence $(P_0, P_1, \cdots, P_\nu)$ of *distinct* points of Γ such that

$$\begin{aligned} P_0 &= a_{j_0} \\ P_1 &= mP_0 + a_{j_1} \\ &\cdots \\ P_\nu &= mP_{\nu-1} + a_{j_\nu} . \end{aligned}$$

Taking into account the m-adic expansion of an integer, we see that the number of elements of Γ which can occur in some sequence σ_ν is precisely $(2m^\nu - 1)^r$ because their coefficients in $\mathbf{Z}$, in terms of our basis, range from $-(m^\nu - 1)$ to $+(m^\nu - 1)$.

Since η is non-degenerate, given $\varepsilon > 0$ there exists an integer $\nu_0 = \nu_0(m, \varepsilon)$ such that for all $\nu \geqq \nu_0$ and any point P_ν in a sequence σ_ν we have

$$(m^d - \varepsilon)\eta(P_{\nu-1}) \leqq \eta(P_\nu) \leqq (m^d + \varepsilon)\eta(P_{\nu-1}) .$$

Inductively, we get

$$(1) \qquad (m^d - \varepsilon)^{\nu-\nu_0}\eta(P_{\nu_0}) \leqq \eta(P_\nu) \leqq (m^d + \varepsilon)^{\nu-\nu_0}\eta(P_{\nu_0}) .$$

Put $c_0(m, \varepsilon) = \log \eta(P_{\nu_0})/\log(m^d + \varepsilon)$.

If $\nu \geqq \nu_0$ and if

$$(2) \qquad \nu \leqq \frac{\log \beta}{\log(m^d + \varepsilon)} + \nu_0 + c_0(m, \varepsilon)$$

then by an obvious logarithm, we see from the right inequality of (1) that

$$\eta(P_\nu) \leqq \beta .$$

Consequently, for such ν, we get a lower bound for $N(\beta)$, namely $(2m^\nu - 1)^r$ which we replace for convenience by $m^{\nu r}/2^r$. We see that for all $\beta \geqq \beta_0(m, \varepsilon)$ we get

$$c_1(m, \varepsilon)\beta^{r\lambda} \leqq N(\beta)$$

where $\lambda = \log m/\log(m^d + \varepsilon)$.

On the other hand, if $\nu \geqq \nu_0$ and

$$(3) \qquad \frac{\log \beta}{\log(m^d - \varepsilon)} - c_0'(m, \varepsilon) + \nu_0$$

with an obvious number $c_0'(m, \varepsilon)$, then by the left inequality of (1) we see that

$$\eta(P_\nu) > \beta .$$

Consequently, all the P with $\eta(P) \leqq \beta$ must lie in sequences σ_ν for ν satisfying

$$(4) \qquad \nu \leqq \frac{\log \beta}{\log(m^d - \varepsilon)} - c_0'(m, \varepsilon) + \nu_0 .$$

This gives us an upper bound for $N(\beta)$ using $2^r m^{\nu r}$ as an upper bound for the number of elements in such sequences, namely,

$$N(\beta) \leqq c_2(m, \varepsilon)\, \beta^{r\lambda'}$$

where $\lambda' = \log m/\log(m^d - \varepsilon)$. Since m, ε where fixed at the beginning of our proof, we can now take m large and ε small, so that the assertion in our theorem is clear.

Historical Note:

The necessity of counting points rational over a number field led Weil to his decomposition theorem [41]. Siegel, in studying integral points, then defined the height. An analysis of all the proofs in this subject shows that the crucial step is always the study of sets of bounded height, as well in the Mordell-Weil theorem as in the Siegel theorem, or the Roth theorem. We have given heights here under a maximally axiomatized situation inspired by Weil's treatment of heights in number fields [40].

With a view to applying this theory to the theorem of the base, Neron [24], [25] studied heights in function fields, where it was again taken up in [19]. The geometric interpretations were already given in [25], especially Theorem 2.

Originally, the height was only considered relatively. It was Northcott [26] who observed that by taking a suitable root, one could get it uniformly for algebraic extensions as well, and could extend to these the theorem, in number fields, that points of bounded degree and bounded height form a finite set. This is not used in our applications, however, since we are always concerned with points rational over the given ground field.

It is interesting to try to give an estimate for the number of elements of a number field K of height $\leqq B$ for $B \to \infty$. Let $r = r_1 + r_2 - 1$ (where r_1 is the number of real absolute values and r_2 the number of complex ones). If one asks for the number of integers or units of height $\leqq B$, then one is again reduced to computing the number of lattice points in a region. For the integers, the volume of this region is of type

$$B(c_0 + c_1(\log B) + \cdots + c_r(\log B)^r)$$

where c_i are constants >0 (combinations of powers of 2 and π). As Cassels pointed out to me, one should then use Davenport's lemma [7] to compute the number of lattice points in this region. One obtains then the asymptotic estimate $c_K B(\log B)^r$ for the number of integers of height $\leqq B$, and similarly, an estimate $c_K(\log B)^r$ for the units of height $\leqq B$.

Over the rationals, there is of course no problem for the integers and units, and it is a standard result of elementary number theory that the number of relatively prime pairs of integers of absolute value $\leqq B$ satisfies an estimate of the type $c_{\mathbf{Q}} B^2$ with an error term $O(B \log B)$.

Gauss' lemma is included here partly for its application to Roth's theorem, where it allows us to start the induction trivially. It is a very interesting question to try, in higher dimensions, to give a relation between the absolute values of the coefficients of a polynomial and the height of the rational points on the variety determined by it.

Proposition 7 is the usual infinite descent of the Mordell-Weil theorem properly axiomatized (cf. the historical note of Chapter *V*). Similarly, Theorem 3 and its proof are taken from Néron's thesis [24].

CHAPTER IV

Properties of Heights

Throughout this chapter, **F** *is a field with a proper set of absolute values* $M_{\mathbf{F}}$ *satisfying the product formula. We denote by* K *some finite extension of* **F**, *so that the set* M_K *satisfies the product formula with multiplicities* N_v, $v \in M_K$.

If V is a variety defined over K, then for each morphism of V into a projective space also defined over K, we can define the absolute height of a point of V algebraic over K, and its relative height if it is rational over K.

We shall assume that V is complete and normal. Maps into projective spaces are then given by linear systems, and it will be our main purpose to give conditions under which the corresponding height functions are equivalent or quasi-equivalent (for the definition, see § 1). We shall formulate our theorems for the absolute height h, and it will then be obvious that they remain valid for the relative height H_K which is a fixed power of the absolute one (namely the $[K:\mathbf{F}]$-th power). In the applications, we need only the equivalence criteria for the relative height.

§ 1. *Functorial properties*

Let V be a variety. Two real-valued functions λ and λ' on a set of points of V will be said to be *equivalent* if there exist two numbers $c_1, c_2 > 0$ such that for all points P in our set, we have

$$c_1\lambda(P) \leqq \lambda'(P) \leqq c_2\lambda(P) .$$

We then write $\lambda \sim \lambda'$. We define λ and λ' to be *quasi-equivalent* (and write $\lambda \approx \lambda'$) if given $\varepsilon > 0$ there exist two numbers $c_1, c_2 > 0$ depending on ε, such that for all points P in our set we have

$$c_1\lambda(P)^{1-\varepsilon} \leqq \lambda'(P) \leqq c_2\lambda(P)^{1+\varepsilon} .$$

These relations are obviously equivalence relations (symmetric, reflexive, transitive).

Let **F** be a field with a proper set of absolute values $M_{\mathbf{F}}$ satisfying the product formula. Then for each finite extension K of **F**, M_K

satisfies the product formula with multiplicities N_v. Let V be a variety defined over K. Let $\varphi\colon V \to \mathbf{P}^m$ be a morphism of V into projective space, defined over K. Then for each point P of V_K, $\varphi(P)$ is a point of $\mathbf{P}^m$, rational over K, and we can thus define its height which we shall denote by $H_{K,\varphi}(P)$. If P is algebraic over $\mathbf{F}$, then there exists a finite extension K of $\mathbf{F}$ over which it is rational, and we can then define its absolute height

$$h_\varphi(P) = h(\varphi(P)) .$$

Thus $H_{K,\varphi}$ is a function on V_K while h_φ is a function on $V_{\bar{K}} = V_{\bar{\mathbf{F}}}$, i.e. on the set of points on V rational over the algebraic closure of K.

Let V be projective, defined over K. Suppose its points are represented by homogeneous coordinates $(x_0, \cdots, x_n)$. Let (a_{ij}) $(i = 0, \cdots, m$ and $j = 0, \cdots, n)$ be a matrix with coefficients in K, and put

$$y_i = a_{i0}x_0 + \cdots + a_{in}x_n .$$

Then the map $(x) \to (y)$ defines a rational map $\varphi\colon V \to \mathbf{P}^m$. If P is a point with coordinates (x) such that not all y_i are equal to 0 in the above formula, then φ is defined at P. A map φ obtained in the manner just described is called a *linear projection*, defined over K.

PROPOSITION 1. *Let V be a projective variety defined over K. Let $\varphi\colon V \to \mathbf{P}^m$ be a linear projection also defined over K. There exists a number $c > 0$ depending only on φ, such that if P is a point of V, algebraic over K, such that not all the coordinates y_i above are 0, then*

$$h_\varphi(P) \leqq ch(P)$$

(h being the height in the given projective embedding of V).

Proof: Let S_K be the subset of M_K containing all those absolute values v for which some $|a_{ij}|_v$ is not 1 and all archimedean absolute values. Then S_K is a finite set. If our coordinates are in a finite extension E of K, then for any $w \in M_E$ extending some $v \in M_K$, we have

$$\sup_i |y_i|_w \leqq c_v \sup_j |x_j|_w$$

where $c_v = (n+1)\sup_{ij}[|a_{ij}|_v, 1]$ for all $v \in S_K$, and $c_v = 1$ for all $v \notin S_K$. From this we get

$$h_\varphi(P) \leqq ch(P)$$

where

$$c = \left[\prod_{v \in M_K} \prod_{w|v} c_v^{N.w} \right]^{1/[E:\mathrm{F}]} \leqq \prod c_v \, .$$

From our proposition, we see in particular that if V is again a complete variety, and if φ, φ' are two morphisms into $\mathbf{P}^m$ which differ by a projective transformation defined over a finite extension K of F, then h_φ is equivalent to $h_{\varphi'}$. In Property 3, we shall prove a much stronger criterion for equivalence.

Let us assume now that V is complete and normal.

We shall obtain mappings $\varphi\colon V \to \mathbf{P}^m$ by means of linear systems (cf. IAG, Chapter VI, or *Foundations*, Chapter IX). Let $\mathscr{L}$ be a linear system on V defined over K. This means that we can find a divisor $X_0 \in \mathscr{L}$ rational over K, such that the space of functions L_0 consisting of all functions on V whose divisors are of type $X - X_0$ as X ranges over $\mathscr{L}$ (together with 0) has a basis defined over K. If $(f_0 = 1, \cdots, f_m)$ is such a basis, then it defines a rational map $\varphi\colon V \to \mathbf{P}^m$. If $\mathscr{L}$ is without fixed points, then φ is a morphism, and conversely if φ is a morphism, the linear system without fixed component belonging to it is without fixed point. In view of the remark in the preceding paragraph, we see that the equivalence class of h_φ actually depends only on the linear system.

Instead of taking a basis as above, let us take any set of generators of the vector space L_0, defined over K, say $g_0, \cdots, g_n$. Then each g_i is a linear combination of the f_j with coefficients in K, and conversely. Our set $(g_0, \cdots, g_n)$ determines a rational map

$$\varphi'\colon V \to \mathbf{P}^n,$$

said to be *derived from* $\mathscr{L}$, and it is clear that φ' is a morphism if and only if φ is a morphism. If this is the case, then we see again that $h_\varphi \sim h_{\varphi'}$.

Let $\mathscr{M}$ be another linear system on V, also defined over K. Then we can define the sum $\mathscr{L} + \mathscr{M}$ as a linear system. If $(f_0 = 1, \cdots, f_m)$ is a basis for L_0 over K, and $(g_0 = 1, \cdots, g_n)$ is a basis for M_0 over K obtained in a similar manner, then we get a vector space of functions N_0 generated by the products $f_i g_j$. (These $f_i g_j$ need not be linearly independent over K, however.) We have $f_0 g_0 = 1$. The divisor of $f_i g_j$ is

$$(f_i g_j) = (f_i) + (g_j) \, .$$

If we write $(f_i) = X_i - X_0$ and $(g_j) = Y_j - Y_0$ with $X_i, X_0 \in \mathscr{L}$ and $Y_j, Y_0 \in \mathscr{M}$ then

$$(f_i g_j) = X_i + Y_j - (X_0 + Y_0) .$$

The vector space N_0 gives rise to a linear system $\mathscr{N}$ called the sum of $\mathscr{L}$ and $\mathscr{M}$. If $\mathscr{L}$ and $\mathscr{M}$ are both without fixed points, so is their sum.

Let $\varphi\colon V \rightarrow \mathbf{P}^m$ and $\varphi'\colon V \rightarrow \mathbf{P}^n$ be the two rational maps derived from $\mathscr{L}$ and $\mathscr{M}$ above. Then the functions $\{f_i g_j\}$ give rise to a rational map ψ of V into $\mathbf{P}^{(m+1)(n+1)-1}$. If we denote by $\varphi + \varphi'$ any one of the rational maps derived from $\mathscr{L} + \mathscr{M}$ then obviously

$$h_\psi = h_\varphi h_{\varphi'} , \qquad \text{and} \quad h_\psi \sim h_{\varphi+\varphi} .$$

Summarizing, we get:

Property 1. *Let V be a complete, abstract, normal variety defined over K. Let $\mathscr{L}$, $\mathscr{M}$ be two linear systems on V, also defined over K, and without fixed points. Let φ, φ' be two rational maps of V into $\mathbf{P}^m, \mathbf{P}^n$ respectively, derived from these systems over K and let $\varphi + \varphi'$ be a rational map derived from $\mathscr{L} + \mathscr{M}$. Then*

$$h_{\varphi+\varphi'} \sim h_\varphi h_{\varphi'} .$$

The next property also follows immediately from the definitions.

Property 2. *Let U, V be two complete abstract varieties defined over K. Assume that U is normal, and V non-singular. Let $\omega\colon U \rightarrow V$ be a morphism, surjective and defined over K. Let $\mathscr{L}$ be a linear system on V defined over K, and let $\omega^{-1}(\mathscr{L})$ be the linear system on U consisting of all divisors $\omega^{-1}(X)$ as X ranges over $\mathscr{L}$. Let φ be a rational map of V into $\mathbf{P}^m$ derived from $\mathscr{L}$ over K. Then $\varphi \circ \omega$ is a rational map derived from the linear system $\omega^{-1}(\mathscr{L})$. Assume in addition that $\mathscr{L}$ is without fixed points. Then so is $\omega^{-1}(\mathscr{L})$ and we have*

$$h_{\varphi\circ\omega} = h_\varphi \circ \omega .$$

Proof: We have assumed that V is non-singular in order to ensure that the inverse image of a divisor linearly equivalent to 0 is also linearly equivalent to 0. That $\omega^{-1}(\mathscr{L})$ is a linear system is obvious, and so is the rest of our assertion.

§ 2. *Criterion for equivalence*

The next property is the most crucial one.

Property 3. *Let V be a complete, abstract, normal variety defined over K. Let $\mathscr{L}$, $\mathscr{M}$ be two linear systems on V, also defined over*

K, and without fixed points. Let φ, ψ be two rational maps of V into $\mathbf{P}^m, \mathbf{P}^n$ respectively, derived from these systems. If the divisors of $\mathscr{L}$ and $\mathscr{M}$ are linearly equivalent to each other, then h_φ and h_ψ are equivalent.

Proof: We shall first need a lemma. In the course of its proof, we shall use the fact that if a function $f \in \bar{K}(V)$ is not defined at a point Q of V, then it has a pole passing through Q (IAG, Proposition 3 of Chapter VI, § 1).

As usual, we let ∞ be greater than any real number. Also, the product formula is irrelevant for the lemma, so we restate the hypotheses completely.

LEMMA 1. *Let K be a field with a proper set of absolute values M_K. Let V be a complete normal variety defined over K. Let $f_1, \cdots, f_m$ be functions in $\bar{K}(V)$ whose divisors of zeros have no point in common. Then there exists an M_K-divisor c such that for all $\bar{K}$-valued places π of $\bar{K}(V)$ over $\bar{K}$ and all $w \in M_{\bar{K}}$ we have*

$$\sup_\alpha | \pi(f_\alpha) |_w \geqq c(w) .$$

Proof: In the ring $K[f_1, \cdots, f_m]$ consider the ideal $(f_1, \cdots, f_m)$. It must be the unit ideal, for otherwise there would exist a homomorphism of our ring over K mapping all the f_α on 0, and this homomorphism can be extended to a place. Following it up by another place if necessary we may assume that this place is $\bar{K}$-valued and that it has been extended to $\bar{K}(V)$. Some conjugate place will then be the identity on $\bar{K}$, and this contradicts our assumption. Hence there exists an expression of the type

$$1 = \sum z_\lambda M_\lambda(f_1, \cdots, f_m)$$

where the z_λ lie in K, and the M_λ are monomials with coefficients 1 and of degree $\geqq 1$. Consider the finite set S_K consisting of all archimedean absolute values and all those v where some $| z_\alpha |_v \neq 1$. We can take $c(v) = 1$ for $v \notin S_K$, and for $v \in S_K$, whenever all the $\pi(f_\alpha)$ are finite, the relation

$$1 = \sum z_\lambda M_\lambda(\pi(f_1), \cdots, \pi(f_m))$$

shows that the $| \pi(f_\alpha) |_w$ cannot all lie below an obvious lower bound, determined by the absolute values of the z_λ. This proves our lemma.

Let X_0 be a divisor of $\mathscr{L}$, rational over K, so that the derived space of functions L_0 has a basis $(f_0, \cdots, f_m)$ defined over K. For each f_i, we can write

$$(f_i) = X_i - X_0 \qquad X_i \in \mathscr{L} .$$

We denote by φ_i the rational map of V into the affine K-open subset of $\mathbf{P}^m$ determined by the functions $(f_0/f_i, \cdots, f_m/f_i)$. We let V_i be the K-open subset $V - \mathrm{Supp}(X_i)$ of V. Then the V_i cover V, and φ_i is defined at every point of V_i.

We proceed similarly with $\mathscr{M}$. In view of our assumption, we may take the divisor $Y_0 = X_0$ to get the derived space of functions M_0 with a basis $(g, \cdots, g_n)$ defined over K. Then

$$(g_j) = Y_j - X_0 , \qquad Y_j \in \mathscr{M} ,$$

and the $(g_0, \cdots, g_n)$ determine our map $\psi\colon V \to \mathbf{P}^n$.

We apply our lemma to the functions $f_0/g, \cdots, f_m/g$ where g is any one of the g_j. Then

$$(f_i/g) = X_i - Y$$

with $Y > 0$ and the hypothesis of the lemma is satisfied. For any point $P \notin \mathrm{supp}(X_0)$, such that $g(P) \neq 0$, we get

$$\sup_i | f_i(P)/g(P) |_w \geqq c(v)$$

for any $w \in M_{\overline{K}}$ extending $v \in M_K$, and hence

$$\sup_i | f_i(P) |_w \geqq c(v) \, | g(P) |_w .$$

This last relation remains valid whether $g(P)$ is 0 or not (always provided $P \notin \mathrm{supp}(X_0)$). Since g was selected arbitrarily among the g_j, we get for any $P \notin \mathrm{supp}(X_0)$

$$\sup_i | f_i(P) |_w \geqq c(v) \sup_j | g_j(P) |_w .$$

If P is rational over the finite extension E of K, then w may be viewed as being in M_E. We raise this inequality to the power $N_w = [E_w \colon \mathbf{F}_w]$, take the product over all $w \in M_E$, and take the $1/[E \colon \mathbf{F}]$-th root. In view of the facts that $c(v) = 1$ for v not lying in some fixed finite subset S_K of M_K, and that our absolute values are proper, we get

$$c_1 h_\psi(P) \leqq h_\varphi(P)$$

where c_1 is the fixed number $\prod_{v \in S_K} c(v)$ which does not depend on the field E, and thus holds for all points P algebraic over K, $P \notin \mathrm{supp}(X_0)$.

Our constant c_1 of course depends on V_0, and we may write $c_1 =$

$c_1(V_0)$. Since V is covered by $V_0, \cdots, V_m$, we can now take for c_1 the smallest of the $c_1(V_i)$, so that our inequality above holds for all P in V algebraic over K.

The inequality on the right follows by symmetry.

§ 3. *Criterion for quasi-equivalence*

From Property 3 and lemmas of pure algebraic geometry, we shall prove that if the divisors of the linear system are algebraically equivalent, then the associated height functions are quasi-equivalent.

LEMMA 2. *Let V be a non-singular projective variety, and X a hyperplane section. Given an integer d, there exists a positive integer e depending only on V and d such that for any positive divisor Y on V of degree d, the divisors $Y + eX$ and $-Y + eX$ are ample.*

Proof: The assertion concerning $-Y + eX$ is proved in F^2-IX_6 Theorem 13. The other assertion is an immediate consequence of this one: namely, we first find $e_1 > 0$ such that $-Y + e_1X$ is ample. It is then linearly equivalent to a positive divisor T, and by cutting with generic linear varieties of suitable dimension, we see that the degree of T depends only on that of Y, and e_1. Repeating our application of Theorem 13, we find an integer $e_2 > 0$ such that $-T + e_2X$ is ample, and e_2 then depends only on d. This proves our lemma.

LEMMA 3. *Let V be a complete non-singular variety. Let X be a divisor on V such that some positive multiple of X is ample. Then there exists an integer $e > 0$ such that for any divisor Z on V algebraically equivalent to 0, the divisor $Z + eX$ is ample.*

Proof: Let $\hat{A}$ be the Picard variety of V. We can always find a Poincaré divisor D on $V \times \hat{A}$ which is positive. If V is an abelian variety, this is AV, Theorem 10 of Chapter IV, § 4, and otherwise, making a generic translation on D, the pull-back method gives a positive Poincaré divisor on $V \times \hat{A}$ (AV, Theorem 1 of Chapter VI, § 1).

By hypothesis, $Z \sim {}^tD(a) - {}^tD(0)$ for some point $a \in \hat{A}$. The intersections are defined after making a generic translation on D. By Lemma 2, there exists an integer $e_1 > 0$ such that $-{}^tD(0) + e_1X$ is ample. Furthermore, the divisors ${}^tD(a)$ as a ranges over $\hat{A}$ are all algebraically equivalent to each other, are positive divisors, and have the same projective degree. Hence there exists an integer $e_2 > 0$ such that ${}^tD(a) + e_2X$ is ample, again by Lemma 2. From this our lemma is immediate.

The following notation will be convenient. If Y is a divisor on V defined over K, such that its complete linear system is without fixed points, then we shall denote by h_Y the height function h_φ associated with any one of the maps φ derived from this linear system. Thus h_Y is well defined, up to equivalence. (Similarly, for points in V_K, we shall write H_Y for H_φ.)

PROPERTY 4. *Let V be a complete non-singular variety defined over K. Let $X, Y > 0$ be two positive divisors on V, rational over K. Assume that a positive multiple of each is ample, and that the linear systems $\mathscr{L}(X)$ and $\mathscr{L}(Y)$ are without fixed points. If X and Y are algebraically equivalent, then h_X and h_Y are quasi-equivalent.*

Proof: Using Property 3, we shall reduce our assertion to a statement concerning linear equivalence of divisors on V. By Lemma 3, there exists an integer $e > 0$ such that for all $n > 0$ we have

$$n(X - Y) + eX \sim Z_n$$

where Z_n is a positive divisor on V, and $\mathscr{L}(Z_n)$ is without fixed points. Since $n(X - Y) + eX$ is rational over K, one may take Z_n rational over K. We get $nX + eX \sim nY + Z_n$ and taking heights, this yields in view of Property 1,

$$h_X^{n+e} \sim h_X^n h_{Z_n} .$$

Since the height is always $\geqq 1$, we take an nth root of our relation for n large. We see that given $\varepsilon > 0$ there exists a number $c > 0$ such that

$$ch_X(P)^{1+\varepsilon} \geqq h_Y(P)$$

for all points P on V, algebraic over K. The other inequality is obtained in a similar way, or by symmetry.

An important case is that in which V is a curve. In that case, algebraic equivalence is determined by the degree, and hence we get:

COROLLARY. *Let V be a curve, i.e. of dimension 1 in the preceding property. If $\deg(X) = d$ and $\deg(Y) = d'$, then h_X is quasi-*

equivalent to $h_Y^{d/d'}$.

One can give a direct proof using the Riemann-Roch theorem formally instead of the Picard variety, as in [37]. Thus when V is of dimension 1, very little machinery is needed to establish Properties 3 and 4.

§ 4. *Projections on curves*

In the application of the theory of heights in Chapter VII we shall deal with curves, and it is therefore convenient to investigate more closely how the height behaves under projection.

Let U be a projective non-singular curve defined over K. The height h is taken with respect to this embedding.

PROPOSITION 2. *Let* $(y_0, \cdots, y_n)$ *with* $y_0 = 1$, *be functions in* $K(U)$ *determining our embedding in* $\mathbf{P}^n$. *Let* Φ *be a finite set of points of* U *in* $\bar{K}$. *Then there is a polynomial equation such that if* $(a_0, \cdots, a_n, b_0, \cdots, b_n)$ *are elements of* K *not satisfying this equation, then the function*

$$y = \frac{a_0y_0 + \cdots + a_ny_n}{b_0y_0 + \cdots + b_ny_n}$$

has the following properties:

(i) *The function* y *is not constant, and has no zero or pole among the points of* Φ.

(ii) *If* h_y *is the height determined by the mapping of* U *into* $\mathbf{P}^1$ *arising from the function* y, *then* $h_y \sim h$.

(iii) *The mapping* $Q \to y(Q)$ *gives an injection of* Φ *into* $\bar{K}$.

Proof: Let Q be a point of U in $\bar{K}$. If t is a function of order 1 at Q, then each y_i has an expansion as a power series in t, say $y_i = \xi_i t^{e_i} + \cdots$ with an integer e_i, which may be negative, and $\xi_i \in \bar{K}$. We see that there is a polynomial G_Q (linear, in fact) such that for any set of elements $(a_0, \cdots, a_n)$ in K for which $G_Q(a) \neq 0$, then $a_0y_0 + \cdots + a_ny_n$ has order e at Q, where $e = \inf_i e_i$. Taking Q from a finite set Φ, we then take the product of the G_Q for $Q \in \Phi$ and achieve the same thing for all $Q \in \Phi$.

If $\mathfrak{a}$ denotes the sup of the polar divisors of our given y_i, then we see that almost all linear combinations $a_0y_0 + \cdots + a_ny_n$ have precisely $\mathfrak{a}$ as polar divisor. Furthermore, applying the above remarks to zeros instead of poles, and taking into account that the linear system determined by $(1, y_1, \cdots, y_n)$ is without a fixed point, we see that we can make a sufficiently general choice of a_i and b_i such that the function y has no zero or pole in Φ, and its divisor

$$(y) = (y)_0 - (y)_\infty$$

is such that $(y)_\infty$ lies in the above linear system. According to Property 3 of heights, it follows that h_y is equivalent to h.

To insure that the map $Q \to y(Q)$ is injective on Φ, we select among the y_i (for each Q) one of the functions having the highest order pole at Q, and denote it by y_Q. All quotients y_i/y_Q are defined at Q, and we have

$$y(Q) = \frac{a_0 w_0(Q) + \cdots + a_n w_n(Q)}{b_0 w_0(Q) + \cdots + b_n w_n(Q)}$$

where $w_i = y_i/y_Q$. (Strictly speaking, each w_i should carry Q also as an index.) We can choose the b_i so that the denominator does not vanish, and the condition that $y(Q) \neq y(Q')$ when $Q \neq Q'$ are two distinct points of Φ is immediately seen to be implied by the nonvanishing of a polynomial in the a_i and b_i. This concludes the proof of our three statements.

Historical Note:

For some time, the properties of heights were mixed up with the properties centering around the decomposition theorem. These originated mainly with Weil, to whom Property 3 is due. Properties 1 and 2 are of course trivial and are stated explicitly only to have a ready formalism available. We have taken our proof of Property 3 from [40].

Property 4 in the case of curves was essentially discovered by Siegel [37], as it was needed for the study of integral points. It was generalized to arbitrary dimensions by Lang [18] (for the same purpose). It can be used efficiently in the infinite descent.

The property of projections has been abstracted from Siegel's and Mahler's papers, always motivated by its application to integral points [18], [21]. It will not be used until Chapter VII.

CHAPTER V

The Mordell-Weil Theorem

We consider abelian varieties, defined over essentially global fields, namely, those of Examples 1, 2, and 3 of Chapter II, § 1. We shall prove an absolute result and a relative one concerning the group of rational points of an abelian variety over such fields, namely:

THEOREM 1. *Let K be a finitely generated field over the prime field. Let A be an abelian variety defined over K. Then A_K is finitely generated.*

To state the second theorem, we recall a definition. Let K be a function field over the constant field k. Let A be an abelian variety defined over K. A *K/k-trace* of A is a pair (B, τ) consisting of an abelian variety B defined over k and a homomorphism

$$\tau: B \to A$$

defined over K having the following universal mapping property: Given an abelian variety C defined over k and a homomorphism $\varphi: C \to A$ defined over K, then there exists a unique homomorphism $\varphi_*: C \to B$ defined over k such that the diagram

$$\begin{tikzcd}
C \arrow[rr, "\varphi"] \arrow[dr, swap, "\varphi*"] & & A \\
& B \arrow[ur, swap, "\tau"] &
\end{tikzcd}$$

is commutative. For the existence of the K/k-trace see AV [17], Theorem 8 of Chapter VIII, § 3. We also recall that if E is an extension of k which is free from K over k, then (B, τ) is also a KE/E-trace: It does not change under constant field extensions, because K is regular over k by assumption. Furthermore, the homomorphism τ is injective.

We can now state our relative result.

THEOREM 2. *Let K be a function field over the constant field k, and let A be an abelian variety defined over K. Let (B, τ) be a K/k-trace of A. Then $A_K/\tau B_k$ is finitely generated.*

For instance, if A is an elliptic curve with transcendental j-invariant over k, and $K = k(j)$, then $B = 0$ and A_K is finitely generated.

We prove the theorem in two steps. First the so-called weak Mordell-Weil theorem, where we prove that for an integer $m \geqq 1$ prime to the characteristic, A_K/mA_K is finite under the following additional assumptions:

The group of points A_m of period m is contained in A_K.

In the function field case, the constant field is algebraically closed, and the dimension of K over k is equal to 1.

Indeed, we prove our result by relating our factor group A_K/mA_K to the Galois group of certain unramified extensions of K. This is the Kummer theory, which we reproduce in § 2. Using the fact that unit groups and ideal class groups are of finite type, we prove that certain unramified extensions are finite, and then that A_K/mA_K is finite. In the function field case, these arguments certainly apply in dimension 1, and reduction steps proved in § 1 show that we can limit ourselves to this case in order to prove Theorem 2. We shall also prove in § 1 that we may assume that $A_m \subset A_K$, and that k is algebraically closed.

Once this is carried out, we apply Property 4 of heights to abelian varieties and see that the logarithm of the height function is of quasi-degree 2 so that we can apply Proposition 7 of Chapter IV, § 4 and its corollary. In number fields, we are then finished. In function fields, we must make an analysis of sets of points of bounded height. We shall find out that they lie in a finite number of cosets of τB_k. This will thus prove Theorem 2.

We then return to Theorem 1 for finitely generated fields over the prime field, viewing the algebraic closure k of the prime field in K as a constant field. Applying Theorem 2, the known result in number fields in characteristic 0, and the trivial fact that B_k is finite if k is finite, we conclude the proof of Theorem 1.

We can then apply Theorems 1 and 2 to get results concerning divisor classes.

§ 1. *Reduction steps*

The propositions which we prove in this setion will be used to reduce our main theorem in function fields to special cases. Throughout this section, K will denote a finitely generated regular extension of a field k.

We first show that to prove our main theorem we may extend our function field by a finite separable extension.

PROPOSITION 1. *Let $L \supset K$ be a finite separable extension of K, also regular over k. Let A be an abelian variety defined over K. Let $(A^{K/k}, \tau_K)$ be its K/k-trace, and $(A^{L/k}, \tau_L)$ its L/k-trace. Then the factor group*

$$[\tau_L(A^{L/k})_k \cap A_K]/\tau_K(A^{K/k})_k$$

is finite.

Proof: Let Γ_L be the graph of τ_L. It is defined over L. Let us take the intersection

$$H = \cap \Gamma_L^\sigma$$

of Γ_L and its conjugates over K, where σ ranges over the distinct isomorphisms of L over k. Then H is K-closed, and is an algebraic subgroup of $A^{L/k} \times A$. Its connected component H_0 it therefore defined over K by Chow's theorem (AV Theorem 5, Chapter II, § 1).

We have

$$\Gamma_L \cap [(A^{L/k})_k \times A_K] = H \cap [(A^{L/k})_k \times A_K]$$

and this group contains $H_0 \cap [(A^{L/k})^k \times A_K]$ as a subgroup of finite index since H_0 is of finite index in H.

We have a surjective homomorphism

$$\Gamma_L \cap [(A^{L/k})_k \times A_K] \to \tau_L(A^{L/k})_k \cap A_K$$

by projection on the second factor. We contend that the inverse image of $\tau_K(A^{K/k})_k$ contains $H_0 \cap [(A^{L/k} \times A_K]$. From this it will be clear that the desired factor group is finite.

Since H_0 is contained in the graph of a homomorphism of $A^{L/k}$ into A, it is itself the graph of a homomorphism β of its projection B on the first factor into A. Furthermore, B is contained in $A^{L/k}$ and is defined over K, and hence over k by Chow's theorem. By the universal mapping property of the K/k-trace there exists a commutative diagram

$$\begin{tikzcd}
B \arrow[rr, "\beta"] \arrow[dr, "\beta_*"'] & & A \\
& A^{K/k} \arrow[ur, "\tau_K"'] &
\end{tikzcd}$$

with β_* defined over k, hence $\beta(B_k)$ is contained in $\tau_K(A^{K/k})_k$. This proves our contention and concludes the proof of our proposition.

Corollary. *If Theorem 2 is true for L/k, then it is true for K/k, i.e. if $A_L/\tau_L(A^{L/k})_k$ is finitely generated, so is $A_K/\tau_K(A^{K/k})_k$.*

Proof: This follows immediately from the proposition and the fact that we have an injection

$$0 \to A_K/[\tau_L(A^{L/k})_k \cap A_K] \to A_L/\tau_L(A^{L/k})_k .$$

Next we show that we may extend the constant field.

Proposition 2. *Let k' be any extension of k which is independent of K over k, and let $K' = Kk'$. Let A be an abelian variety defined over K and let (B, τ) be its K/k-trace. Then*

$$A_K \cap \tau B_{k'} = \tau B_k .$$

Proof: We know that (B, τ) is also a K'/k'-trace of A. The inclusion $\tau B_k \subset A_K \cap \tau\beta_{k'}$ is obvious. Conversely, let b be a point of $B_{k'}$ such that τb is rational over K. If we knew that τ is regular, i.e. that it is birational between B and its image in A, then in view of the fact that τ is defined over K, we would see immediately that $\tau^{-1}\tau b$ is rational over k. As we do not know that τ is regular, we use Chow's regularity theorem of AV Chapter 9. We can write $K = k(u)$ for suitable parameters u, and $A = A_u$, $\tau = \tau_u$. Let $u_1, \cdots, u_m$ be independent generic specializations of u over k'. For m large, the map

$$x \to (\tau_{u_1}x, \cdots, \tau_{u_m}x)$$

of B into $A_{u_1} \times A_{u_2} \times \cdots \times A_{u_m}$ is regular, and thus birational between B and its image. If τb is rational over K, then $(\tau_{u_1}b, \cdots, \tau_{u_m}b)$ is rational over $k(u_1, \cdots, u_m)$; hence, b is rational over

$$k(u_1, \cdots, u_m).$$

Since we took $k(u_1, \cdots, u_m)$ free from k' over k, and since b is rational over k', it follows that b is rational over k. This concludes the proof.

Finally, we give the reduction to the case where K is of dimension 1 over k.

Proposition 3. *Let $K \supset E \supset k$ be a tower of fields such that K is regular over E and E regular over k. If Theorem 2 is true for K/E and E/k, then it is true for K/k.*

Proof: Let A be an abelian variety defined over K, and let $(A^{K/E}, \tau_{K/E})$ be a K/E-trace of A. Let $(A^{E/k}, \tau_{E/k})$ be an E/k-trace of $A^{K/E}$. By the universal property, there is an injective homomorphism $\beta\colon A^{E/k} \to A^{K/k}$ defined over k such that the following diagram is commutative:

$$\begin{tikzcd}
A & \\
A^{K/E} \arrow[u, "\tau_{K/E}"] & A^{K/k} \arrow[ul, "\tau_{K/k}"'] \\
A^{E/k} \arrow[u, "\tau_{E/k}"] \arrow[ur, "\beta"'] &
\end{tikzcd}$$

All the homomorphisms are injective. We have

$$A_K \supset \tau_{K/E}(A^{K/E})_E \supset \tau_{K/E}\tau_{E/k}(A^{E/k})_k = \tau_{K/k}\beta(A^{E/k})_k \, .$$

If we assume that $(A^{K/E})_E$ is of finite type modulo $\tau_{E/k}(A^{E/k})_k$, it follows that A_K modulo $\tau_{K/k}\beta(A^{E/k})_k$ is also of finite type. But this last group is contained in $\tau_{K/k}(A^{K/k})_k$. We see, therefore, that it suffices to prove Theorem 2 for each step K/E and E/k of the tower.

It is well known and easy to show that one can always construct a tower $E_0 = k \subset E_1 \subset \cdots \subset E_n = K$ such that each step E_i/E_{i-1} is regular and of transcendence degree 1, Such a construction comes from the lemma of Bertini's theorem, which states that if x, y are two elements of K, algebraically independent over k, and such that $y \notin K^p k$, then for all but a finite number of contants $c \in k$, K is regular over $k(x + cy)$. (Cf. IAG, Chapter VI.) If k is finite, one needs a small additional argument. However, in view of Proposition 2, for our purposes, we may assume that k is infinite.

The above reduction steps have thus reduced the proof of Theorem 2 *to the case where* A_K *contains* A_m, k *is algebraically closed, and* K *is of transcendence degree* 1 *over* k.

§ 2. *Kummer theory*

Throughout this section, we let K *be a field and* m *an integer* $\geqq 1$, *prime to the characteristic of* K.

An abelian extension E of K will be said to be of *exponent* m if $\sigma^m = 1$ for every automorphism σ of E over K.

Let A be a commutative group variety defined over K. As usual, we denote its group of rational points in K by A_K. We let A_m be the subgroup of those points $a \in A$ such that $ma = 0$ (writing A additively).

We recall that the map $x \to mx$ is surjective and separable (so that in particular, A_m is finite, of order equal to $\nu(m\delta)$). Quick proof: If $f_1, \cdots, f_r$ are local paramenters at the origin, then

$$f_i(mP) \equiv m \cdot f_i(P) \bmod \mathfrak{m}^2$$

where $\mathfrak{m}$ is the maximal ideal of the local ring at the origin (IAG, Proposition 1 of Chapter IX, § 1). Thus $m\delta$ is generically surjective and separable. If a is a given point, and x a generic point over $K(a)$, then $a - x$ is generic over K. By an m-th *root* of a we shall mean a point b such that $mb = a$. Any two such m-th roots differ by a point of A_m. If $my = x$ and $mz = a - x$ then $m(y + z) = a$. Since $K(x, y)$ and $K(x, z)$ are separable over $K(x)$ and over K, it follows that $K(y + z)$ is also separable over K.

If $a \in A_K$, we denote by $K(1/m \cdot a)$ the field obtained by taking the composite of all fields $K(b)$ where b ranges over all m-th roots of a. Then all points of A_m are rational over that field, which is separable over K by what we have just proved. If $A_m \subset A_K$, then this field is equal to $K(b)$ where b is any one of the m-th roots of a.

From now on, we assume that A_m is contained in A_K. Let B be a subgroup of A_K containing mA_K (i.e. the set of all points ma with $a \in A_K$). We denote by

$$E_B = K(1/m \cdot B)$$

the field obtained by taking the composite of all fields $K(1/m \cdot a)$ where a ranges over B. It is determined by B, in a given algebraic closure of K, as usual.

If σ is any isomorphism of E_B over K, then with the above notation and with $m\alpha = a$, we have

$$m(\sigma\alpha) = \sigma(m\alpha) = m\alpha \; .$$

Consequently $\sigma\alpha - \alpha$ is an element of A_m, say t_σ. Thus

$$\sigma\alpha = \alpha + t_\sigma \; .$$

Hence σ is in fact an automorphism of E_B over K, and E_B is therefore Galois. If τ is another automorphism of E_B over K, and $\tau\alpha = \alpha + t_\tau$ then clearly $\sigma\tau\alpha = \alpha + t_\sigma + t_\tau$ and hence E_B is abelian.

Let G be its Galois group. We define a bilinear map

$$G \times B \to A_m$$

as follows: Let $\sigma \in G$ and $b \in B$. Take any $\beta \in A$ such that $m\beta = b$. Then $\sigma\beta - \beta$ is clearly independent of our choice of β, and is an element of A_m, which we denote by (σ, b). The reader will immediately verify that this is bilinear in σ and b.

By definition, the kernel on the left consists of those σ such that $(\sigma, b) = 1$ for all $b \in B$. As E_B is generated by elements of type $1/m \cdot b$ any such σ must be the identity.

The kernel on the right consists of those b such that $(\sigma, b) = 1$ for all $\sigma \in G$. For any such b, the extension $K(1/m \cdot b)$ is fixed under G and hence must be K itself. This means that $1/m \cdot b$ is in fact rational over K, i.e. that b lies in mA_K. We thus get a bilinear map

$$G \times B/mA_K \to A_m$$

whose kernel on both sides is equal to the identity. Since A_m is finite, we obtain

Proposition 4. *Let A be a commutative group variety defined over the field K. Let m be an integer $\geqq 1$, prime to the characteristic. Let $A_m \subset A_K$, and let B be a subgroup of A_K containing mA_K. Then $K(1/m \cdot B)$ is an abelian extension of exponent m and is a finite extension if and only if B/mA_K is a finite group.*

Let us consider in greater detail the special case where A is the multiplicative group. We then write the group law multiplicatively. Our group A_m is K^*_m, the group of m-th roots of unity, and is cyclic of order m. Thus the two groups G and B/K^{*m} are dual to each other under our pairing above, and of the same order if they are finite. (We shall not need the infinite case in the sequel.)

If B_1, B_2 are two subgroups of K^* containing K^{*m} and if $B_1 \subset B_2$ then $K(B_1^{1/m}) \subset K(B_2^{1/m})$. Conversely, assume this is the case. We wish to prove $B_1 \subset B_2$. Let $b \in B_1$. Then $K(b^{1/m}) \subset K(B_2^{1/m})$ and $K(b^{1/m})$ is contained in a finitely generated subextension of $K(B_2^{1/m})$. Thus we may assume without loss of generality that B_2/K^{*m} is finitely generated. Let B_3 be the subgroup of K^* generated by B_2 and b. Then $K(B_2^{1/m}) = K(B_3^{1/m})$ and from what we have see above, the degree of this field over K is precisely

$$(B_2 : K^{*m}) \text{ or } (B_3 : K^{*m}).$$

Thus these two indices are equal, and $B_2 = B_3$. Summarizing:

Theorem 3. *Let K be a field, m an integer $\geqq 1$ prime to its characteristic. Assume that the* m-th *roots of unity lie in K. Then there is a bijective correspondence between abelian extensions E of K of exponent m and subgroups B of K^* containing K^{*m}, given by*

$E = K(B^{1/m})$, and

$$[E:K] = (B:K^{*m}),$$

in the sense that if one index is finite, so is the other, and they are equal.

Proof: We have proved everything except that every abelian extension of K of exponent m is of type $K(B^{1/m})$ for some B. But such an extension is a composite of finite cyclic extensions of exponents m. Our assertion follows then from Galois theory (Hilbert's theorem 90).

We shall say that the abelian extension $K(B^{1/m})$ and B *belong* to each other. Theorem 3 is said to express the Kummer theory, and extensions of this type are called *Kummer extensions*.

We wish now to determine when a Kummer extension is unramified for a discrete valuation.

PROPOSITION 5. *Let v be a discrete absolute value on K, and $a \in K^*$. Assume that m is prime to the characteristic of the residue class fiield of v. Then v is unramified in $K(a^{1/m})$ if and only if* $\mathrm{ord}_v a$ *is divisible by m.*

Proof: Assume first that the order of a is divisible by m. Then we can write $a = \pi^{rm}u$ where π is a prime element, r an integer, and u a unit. We see that $K(a^{1/m}) = K(u^{1/m})$ and that $u^{1/m}$ is a unit, satisfying $X^m - u = 0$. Since m is assumed to be prime to the characteristic, we can apply the criterion of Proposition 10 of Chapter I, § 3 to conclude that $K(a^{1/m})$ is unramified at any extension of v. Conversely, assume this the case. Then if $a = \pi^s u$ with some integer s, it must be so that m divides s, otherwise, the value group of our extension $K(a^{1/m})$ would obviously be unequal to the value group $v(K^*)$, and so the extension could not be unramified.

Let K have a proper set of absolute values M_K and assume that all the non-archimedean absolute values of M_K are discrete. We shall say that an algebraic extension E of K is *unramified* over M_K, or that M_K is *unramified* in E, if every non-archimedean $v \in M_K$ is unramified in E.

We continue to assume that the m-th roots of unity lie in K, and we let E be a Kummer extension of K, of exponent m, belonging to the subgroup B of K^*. Using Proposition 2, we see that M_K is unramified in E if and only if, for each $b \in B$, there exists an M_K-ideal $\mathfrak{b}$ such that

$$(b) = m\mathfrak{b} .$$

Hence:

THEOREM 4. *Let K be a field with a proper set of absolute values M_K. Assume all the non-archimedean ones are discrete. Let m be an integer prime to the characteristics of the residue class fields of the discrete $v \in M_K$, and assume all m-th roots of unity lie in K. Let B_u be the subgroup of K^* belonging to the maximal abelian extension of K, of exponent m, and unramified over M_K. Then B_u consists of those elements $b \in K^*$ for which there exists an M_K-ideal $\mathfrak{b}$ such that $(b) = m\mathfrak{b}$. If C is the group of M_K-ideal classes, C_m those of period m, and U the group of M_K-units, then we have an exact sequence*

$$0 \to U/U^m \to B_u/K^{*m} \to C_m \to 0 .$$

Proof: If $b \in B_u$, and we have $(b) = m\mathfrak{b}$ as above, then we map b on the ideal class of $\mathfrak{b}$, which is obviously in C_m. The map is surjective in view of Proposition 5, and K^{*m} is obviously contained in the kernel. Thus we get the map of B_u/K^{*m} onto C_m. The reader will immediately verify that its kernel consists of the units modulo $U^m = U \cap K^{*m}$.

COROLLARY. *If in the above theorem, both C_m and U/U^m are finite, then so is the maximal abelian extension of K, of exponent m, and unramified over M_K.*

One can often reduce the study of ramified extensions to that of unramified extensions, over a larger ground field, as follows:

LEMMA 1. *Let K be a field, m an integer $\geqq 1$ prime to its characteristic, and E and abelian extension of exponent m. Let v be a discrete valuation on K, and π a prime element. Assume that the m-th roots of unity lie in K and that m is prime to the characteristic of the residue class field of v. Let $L = K(\pi^{1/m})$. Then every absolute value w of L extending v is unramified in EL.*

Proof: This is immediate from Proposition 5.

We recall that an unramified extension stays unramified under translation (Proposition 11 of Chapter I, § 3). Thus repeated application of Lemma 1 kills the ramification at any given finite number of absolute values of K, by means of a finite extension. We apply this to get the fundamental result:

THEOREM 5. *Let K be a number field, or a function field in one variable over the algebraically closed constant field k. Let m be an integer $\geqq 1$, and prime to the characteristic. Let M_K be the canonical set of absolute values of K, as in Examples* 1 *and* 2 *of Chapter* II, § 1, *and let S be a finite subset of M_K. Then there is only a finite number of abelian extension of K, of exponent m, which are unramified over $M_K - S$.*

Proof: We may assume S contains the divisors of m. Let E be the composite of all abelian extension of K, of exponent m, unramified over $M_K - S$. Let L be obtained by adjoining the m-th roots of unity to K, and also $\pi_1^{1/m}, \cdots, \pi_s^{1/m}$ where $\pi_1, \cdots, \pi_s$ are prime elements at the discrete valuations of S. Then EL, viewed as an extension of L, is unramified over M_L by our lemma. We now apply the corollary of Theorem 4, the finiteness of the class number of Theorem 1, Chapter II, § 2, and the unit theorem in number fields. In the function field case, the units are merely the non-zero constants, and the existence of the Jacobian guarantees that the divisor classes of exponent m form a finite group. Thus our corollary also works in this case.

§ 3. *The weak Mordell-Weil theorem*

If A is an abelian variety defined over a field K, and v is a discrete valuation on K, then we can define the reduction of A with respect to this valution [36]. We denote it by A'_v, and use the prime to denote reduction of objects like cycles, points, etc. In general, A'_v is merely a cycle.

If A'_v has one component, with multiplicity 1, which is an abelian variety whose law of composition is obtained by reducing that of A, then we shall say that the reduction is *non-degenerate*. The following lemma can be proved by a systematic procedure which follows the one at the end of [36].

LEMMA 2. *Let K be a field with a proper set M_K of discrete valuations. Let A be an abelian variety defined over K. Then for all but a finite number of $v \in M_K$, the reduction A'_v is non-degenerate.*

Proof: We leave it to the reader. It consists in looking separately at each geometric property entering into the definition of an abelian variety and its reduction: The fact that the variety remains non-singular, that certain maps remain morphisms, that the associative law holds, etc. In each case, one proves that there exists an element $a \in K^*$ such that if $|a|_v = 1$, then the reduction A'_v has the corresponding geometric property.

Lemma 3. *Let K be a field with a discrete valuation v. Let A be an abelian variety defined over K, and having a non-degenerate reduction A'_v. Let m be an integer $\geqq 1$, prime to the characteristic of K and to the characteristic of the residue class field of v. Assume $A_m \subset A_K$. Then v is unramified in the extension $K(1/m \cdot A_K)$.*

Proof: Since our statement is local in v, we may assume that K is complete under our valuation. Furthermore, it suffices to prove that for each $a \in A_K$, the extension $K(1/m \cdot a)$ is unramified over K. Let P be a point on A such that $mP = a$. The automorphisms of $K(P)$ over K are given by translations as in § 1, and if

$$n = [K(P) : K]$$

then there are n elements $a_1, \cdots, a_n$ of A_m such that the automorphisms σ_i $(i = 1, \cdots, n)$ satisfying

$$\sigma_i P = P + a_i$$

give all automorphisms of $K(P)$ over K. Let a'_i be the reduced points. Since m is prime to the characteristic of the residue class field, it follows that $a'_1, \cdots, a'_n$ are distinct. If φ is any place of $K(P)$ extending the canonical place of v, then $\varphi\sigma_i$ are also places of $K(P)$ and

$$\varphi\sigma_i(P) = (P + a_i)' = P' + a'_i \, .$$

Hence the $\varphi\sigma_i$ are distinct, and by Proposition 8 of Chapter I, § 3 we conclude that $K(P)$ is unramified over K, as desired.

From the two lemmas, Proposition 4 of § 2, and Theorem 4 of § 2, we get:

Proposition 5. *Let K be a number field, or a function field in one variable over an algebraically closed constant field k. Let m be an integer $\geqq 1$, prime to the characteristic of K, and A an abelian variety defined over K such that $A_m \subset A_K$. Then A_K/mA_K is finite.*

Of course, in the function field case, we shall later eliminate the restriction that the function field is of dimension 1. For the moment, we turn to the full Mordell-Weil theorem.

§ 4. *The infinite descent*

Let $\mathbf{F}$ be a field with a proper set of absolute values $M_{\mathbf{F}}$ satisfying the product formula, and let K be a finite extension of

F, with its set of absolute values M_K. Let A be an abelian variety defined over K, which we assume to be embedded in projective space by means of the complete linear system $\mathscr{L}$. Let $X \in \mathscr{L}$, let m be an integer > 1, and $a \in A_K$. Let

$$\omega : A \to A$$

be the isogeny $\omega u = mu + a$. Then

$$\omega^{-1}(X) = (m\delta)^{-1}X_{-a} \equiv m^2 X_{-a} \equiv m^2 X$$

where $\equiv$ is algebraic equivalence (cf. AV, Proposition 2 of Chapter IV, § 1 together with the known fact that the torsion group

$$D_\tau(A)/D_a(A)$$

is trivial). By properties 1, 2, and 4 of heights, we see that $H \circ \omega$ is quasi-equivalent to H^{m^2}, if we denote by $H = H_{M_K}$ the height function determined by our given projective embedding and our set M_K. *In particular, if η is the logarithm of the height, then it defines a function of quasi-degree* 2 *on A_K in the sense of Chapter* III, § 4 and we can apply Proposition 7 of that chapter. If K is a number field, then η is non-degenerate by Theorem 1 of Chapter III, § 2 and consequently we have proved Theorem 1 in this case.

Furthermore, Theorem 3 of Chapter III, § 4 becomes applicable, and thus we have proved:

THEOREM 5. *Let A be an abelian variety defined over a number field K, and in projective space. Then A_K is finitely generated. Given $\varepsilon > 0$, there exist two numbers $c_1(\varepsilon)$, $c_2(\varepsilon) > 0$ and a number $\delta, 0 < \delta < \varepsilon$, such that for all $B > B_0(\varepsilon)$ the number of points $P \in A_K$ with $H(P) \leqq B$ satisfies the inequalities*

$$c_1(\varepsilon)(\log B)^{r/2-\delta} \leqq N(B) \leqq c_2(\varepsilon)(\log B)^{r/2+\delta}$$

where r is the rank of A_K.

To conclude our proof of Theorem 2, we must investigate the nature of sets of points of bounded height.

§ 5. *Points of bounded height*

Let $W \subset \mathbf{P}^r$ be a projective variety, non-singular in codimension 1, defined over a field k. Let w be a generic point of W over k, and $K = k(w)$ a function field for W over k. Let $A = A_w$ be an abelian variety defined over K, and embedded in projective space $\mathbf{P}^n$ over K. Let u be a generic point of A over K. We shall denote by $W \circ A$ the locus of (w, u) over k. Then $W \circ A$ is the graph of the algebraic family parametrized by W, of which A is a generic member. If we put $Z = W \circ A$, then

$$A = Z(w) = \mathrm{pr}_2\ [Z\cdot(w \times \mathbf{P}^n)].$$

Given a rational point P of A over K, we shall denote by W_P the locus of (w, P) over k. Then

$$W_P \subset W \circ A \subset W \times \mathbf{P}^n \subset \mathbf{P}^r \times \mathbf{P}^n .$$

Note that W_P and W are birationally equivalent under the projection pr_1.

There is a projective embedding

$$\varphi: \mathbf{P}^r \times \mathbf{P}^n \to \mathbf{P}^{(r+1)(n+1)-1}$$

which to each product of points with homogeneous coordinates $(x_0, \cdots, x_r) \times (y_0, \cdots, y_n)$ assigns the point with homogeneous coordinates $(x_i y_j)$. One verifies immediately that for any $P \in A_K$ we have

$$H_\varphi(w, P) \leqq H(w) + H(P) = \deg W + H(P) .$$

Here, of course, $H = H_{M_K}$ is the height, taken relative to our projective embeddings and the canonical set of absolute values on K obtained from the prime rational divisors of W over k.

Let c_3 be a number > 0 and S a subset of A_K such that $H(P) \leqq c_3$ for all $P \in S$. Then by Theorem 2 of Chapter III, § 3 there exists a number c_4 such that $\deg \varphi(W_P) \leqq c_4$ and hence, by the theory of Chow coordinates, W_P can belong only to a finite number of algebraic families on $W \circ A$ for $P \in S$. We shall be through with the proof of Theorem 2 once we have proved:

THEOREM 6. *Let W be a projective variety, non-singular in codimension* 1, *defined over a field k. Let w be a generic point of W over k, and $K = k(w)$ a function field of W over k. Let A be an abelian variety defined over K, and (B, τ) be its K/k-trace. For each point $P \in A_K$, let W_P be the locus of (w, P) over k. If $P, P' \in A_K$ are such that W_P and $W_{P'}$ are in the same algebraic family on $W \circ A$, then P and P' are congruent modulo τB_k* (i.e. $P - P' \in \tau B_k$).

Proof: In view of Proposition 2 of § 1, we may assume that k is algebraically closed. We shall prove our theorem by going from W_P to $W_{P'}$ passing through a generic member of the family. Our first step is to show that a generic member is of the same type as W_P.

LEMMA. 3. *Let U be a subvariety of $W \circ A$, defined over an extension k_1' of k independent of K over k, and such that W_P is a specialization of U over k. Put $K_1' = Kk_1'$. Then there exists a rational point P_1 of A in K_1 such that $U = W_{P_1}$.*

Proof: We have $\mathrm{pr}_1 W_P = W$. Hence by the compatibility of projections with specializations, we must have $\mathrm{pr}_1 U = W$. But dim $U =$ dim W, and hence a generic point of U over k_1 is of type (w, Q) for some $Q \in \mathbf{P}^n$, algebraic over $k_1(w)$. We must show that Q is rational over $k_1(w)$, and that Q is in A.

On $W \times \mathbf{P}^n$ the intersection $U \cdot (w \times \mathbf{P}^n)$ is obviously defined, and in view of the above remarks, we have

$$U \cdot (w \times P^n) = \sum_{i=1}^{d} (w, Q_i)$$

for suitable points Q_i. Taking the projection on W, we see that $d = 1$. Since $U \cdot (w \times \mathbf{P}^n)$ is rational over $k_1(w)$, so is $Q_1 = P_1$. Finally, we have $P_1 \in A$ because

$$w \times P_1 = U \cdot (w \times \mathbf{P}^n) \subset (W \circ A) \cdot (w \times \mathbf{P}^n) = w \times A\,.$$

Note that in the lemma, we have made no use of the fact that A is an abelian variety: The same statement holds for any algebraic system of varieties.

It will suffice to prove Theorem 6 for W_P and W_{P_1}. Indeed, given W_P and $W_{P'}$ as in the theorem, there exists a subvariety W_{P_1} of $W \circ A$, defined over some k_1, such that both W_P and $W_{P'}$ are specializations of W_{P_1}. If $P - P_1 \in \tau B_{k_1}$ and $P' - P_1 \in \tau B_{k_1}$, then $P - P' \in \tau B_{k_1}$, and by Proposition 2 of §1 we get $P - P' \tau B_k$.

The field k_1 may be assumed to be of transcendence degree 1 over k. Indeed, there exists a variety T, a generic point t_1, and a point t of T such that W_P is the unique specialization of W_{P_1} over the specialization $t_1 \to t$ (over k). The specialization $t_1 \to t$ over k can be extended to a place of $k(t_1)$ corresponding to a discrete valuation ring by blowing up our point t. Thus there exists a field k_1 such that $k \subset k_1 \subset k(t_1)$, $k(t_1)$ has dimension 1 over k_1, and the specialization $t_1 \to t$ is in fact over k_1. Using Proposition 2 of §1 we may replace k by k_1, and in fact may go over to the algebraic closure of our constant field, without loss of generality for the proof of Theorem 6, and we may assume that T is nonsingular.

We have the following diagram of fields:

$$\begin{tikzcd}
 & K_1 = k_1(w) \arrow[dl, dash] \arrow[dr, dash] & \\
K = k(w) \arrow[dr, dash] & & k_1 = k(t_1) \arrow[dl, dash] \\
 & k &
\end{tikzcd}$$

Let $\mathrm{Alb}(T)$ be the Albanese variety of T, defined over k, and let $f\colon T \to \mathrm{Alb}(T)$ be a canonical map, defined over k. Let

$$g\colon T \to A$$

be the rational map defined over K by the expression $g(t_1) = P_1$. Then there is a homomorphism $\alpha\colon \mathrm{Alb}(T) \to A$ and a point $a \in A$ such that the following diagram is commutative:

$$\begin{tikzcd}
T \arrow[r, "f"] \arrow[d, "g"'] & \mathrm{Alb}(T) \arrow[dl, "\alpha + a"] \\
A &
\end{tikzcd}$$

By the definition of the K/k-trace, there exists a homomorphism $\beta\colon \mathrm{Alb}(T) \to B$ defined over k such that the following diagram is commutative:

$$\begin{tikzcd}
 & \mathrm{Alb}(T) \arrow[dl, "\alpha"'] \arrow[d, "\beta"] \\
A & B \arrow[l, "\tau"]
\end{tikzcd}$$

Since t is simple on T, f is defined at t. Furthermore, since W_P is the unique specialization of W_{P_1} over $t_1 \to t$ (relative to k or K), it follows that P is the unique specialization of P_1 over $t_1 \to t$ (relative to K). Hence $g(t) = P$, and

$$\begin{aligned} P_1 - P = g(t_1) - g(t) &= \alpha f(t_1) + a - \alpha f(t) - a \\ &= \tau\beta f(t_1) - \tau\beta(t) \\ &= \tau b \end{aligned}$$

with $b = \beta[f(t_1) - f(t)] \in B_{k_1}$. This proves our theorem.

Corollary. *Let A be projectively embedded in the above theorem, and let the height be taken relative to this embedding and the canonical set of absolute values $M_{k(W)}$. Let S be a subset of A_K of* *bounded height. Then the points of S lie in a finite number of cosets of* τB_k.

Proof: As we saw in our discussion preceding the theorem, points P of bounded height are such that the varieties W_P have bounded degree, and thus we can apply the theorem.

§ 6. *Theorem of the base*

Let V be a projective variety, non-singular in codimension 1, and defined over the algebraically closed field k which we may take to be a universal domain. The group of divisors $D(V)_k$ on V, rational over k contains the usual subgroups of divisors which are algebraically equivalent to 0 and linearly equivalent to 0, respectively.

$$D(V)_k \supset D_a(V)_k \supset D_l(V)_k \, .$$

The Picard group $D_a(V)_k/D_l(V)_k$ has the structure of an abelian variety, the Picard variety. The *theorem of the base* states:

THEOREM 7. *Let* V, k *be as above. Then the group* $D(V)_k$ *modulo* $D_a(V)_k$ *is of finite type.*

We shall show below how this theorem can be reduced to Theorem 2. First, let us prove a corollary.

COROLLARY. *Let K be a field of finite type over the prime field, and let V be a projective variety, non-singular in codimension* 1, *and defined over K. Let* $D(V)_K$ *be the group of divisors on V rational over K, and* $D_l(V)_K$ *those which are linearly equivalent to zero and rational over K. Then the factor group*

$$D(V)_K/D_l(V)_K$$

is of finite type.

Proof: To begin with, we may extend K by a finite extension so that V has a rational point. Then $D_a(V)_K/D_l(V)_K$ is represented by the group $\hat{A}_K$ of the rational points on the Picard variety, over K. If k is the algebraic closure of the prime field in K, then we just apply Theorem 2 and the known theorem over number fields to conclude that $\hat{A}_K$ is finitely generated. We can now use the theorem of the base to finish the proof.

Let us now prove Theorem 7. Let L_u by a generic linear variety over k, such that $V \cdot L_u = C_u$ is a generic curve (cf. IAG, Chapter 7). Let J be the Jacobian of C_u, defined over the field $k(u)$. Let

$D_0(V)_k$ be the subgroup of $D(V)_k$ consisting of those divisors $X \in D(V)_k$ such that $X \cdot C_u$ is of degree 0. Then $D(V)_k / D_0(V)_k$ is infinite cyclic, and it will suffice to prove that $D_0(V)_k / D_a(V)_k$ is of finite type.

Let $\varphi \colon Z_0(C_u) \to Z_0(J)$ be the canonical map of the 0-cycles of degree 0 on C_u into the 0-cycles on the Jacobian. Then

$$\mathfrak{a} \to S(\varphi(\mathfrak{a}))$$

maps $Z_0(C_u)$ into J, and preserves rationality (i.e. if $\mathfrak{a}$ is rational over $k(u)$, so is $S(\varphi(\mathfrak{a}))$). We get a homomorphism

$$\lambda \colon D_0(V)_k \to J_{k(u)}$$

given by the formula

$$\lambda(X) = S(\varphi(X \cdot C_u)) .$$

According to an elementary equivalence criterion [43], one knows that the kernel E of λ, which contains $D_l(V)_k$, is of finite type modulo $D_l(V)_k$. One can in fact prove that this kernel is precisely $D_l(V)_k$ if we have chosen our projective embedding suitably. This is irrelevant here, and we see in any case that $[E + D_l(V)_k]/D_l(V)_k$ is of finite type. To prove the theorem of the base, it will therefore suffice to prove that $D_0(V_k)/[E + D_a(V)_k]$ is of finite type.

The inverse image $\lambda^{-1}(\lambda(D_a(V)_k))$ is precisely equal to

$$E + D_a(V)_k .$$

We have therefore an injection

$$0 \to D_0(V)_k / [E + D_a(V)_k] \to J_{k(u)} / \lambda D_a(V)_k .$$

Now let $\psi \colon V \to A$ be a canonical map of V into its Albanese variety, defined over k. For a suitable constant $b \in A$, we have a commutative diagram

$$\begin{tikzcd}
C_u \arrow[r, "i"] \arrow[d, "\varphi"'] & V \arrow[d, "\psi"] \\
J \arrow[r, "i_* + b"'] & A
\end{tikzcd}$$

where i_* is induced by the inclusion.

Let us look at the inverse images of divisors in $D_a(V)_k$ under the composite maps $(i_* + b) \circ \varphi$ and $\psi \circ i$. The formalism $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$ can be applied to the first according to Appendix 1 of AV since J and A are non-singular. A direct verification using intersection theory (associativity and the definition of $C_u = V \cdot L_u$) shows that it can also be applied to the second. According to the definition of the Picard variety, which we know is obtained by pull back from that of the Albanese variety, we see that the map

$$\lambda \colon D_a(V)_k \to J_k(u)$$

induces the rational homomorphism ${}^t i_*$ on $\hat{A}_k = D_a(V)_k/D_l(V)_k$ if we denote as usual by the upper index t the transpose of i_* on the Picard varieties. Consequently we have an injection

$$0 \to D_0(V)_k/[E + D_a(V)_k] \to J_{k(u)}/{}^t i_* \hat{A}_k \, .$$

By Chow's theory of the $k(u)/k$-trace, one knows that $(\hat{A}, {}^t i_*)$ is a $k(u)/k$-trace of $\hat{J} = J$ (AV, Chapter 8, Theorem 12). Consequently we can now apply Theorem 2, and the theorem is proved.

Historical Note:

It was Mordell [23] who discovered that the group of rational points (over $\mathbf{Q}$) of an elliptic curve is finitely generated. (It is also intersting to recall that in this paper, he conjectures that the set of integral points is finite, and that the set of rational points on a curve of genus $\geqq 2$ is finite. This second conjecture is still unproved.)

The extension to number fields and abelian varieties offered serious problems, because at that time, among other things, it was not clear how to estimate the size of a point. In Mordell's case, over the rationals, there was no problem. For this purpose, Weil invented his decomposition theorem and the theory of distributions. Actually, as was shown by Lang-Néron [19], one does not need the full theory to carry out the infinite descent: The definition of heights and the standard properties suffice. However, the principle of the proof has remained the same since the early days of the theory, first the finiteness of A_K/mA_K, and second the infinite descent. As was said above, this was done by Weil in his thesis [41].

I have given the proof of Lang-Tate for the first. Another method will be found in Roquette [29], but one should note that the finite generation of the unit group and divisor class group and the nondegeneracy of a reduction for almost all $\mathfrak{p}$ are the essential ingredients of these proofs.

As the techniques of algebraic geometry became better, and properties or heights more organized, the infinite descent could be somewhat slicked up (cf. the progression [41], [24], [19]).

Some time ago, Severi noticed the analogy between the theorem of the base on surfaces and the Mordell-Weil theorem. Making the analogy precise, Néron in his thesis [24] proved Theorem 1 (i.e. the Mordell-Weil theorem for fields of finite type) and also the theorem of the base, in any characteristic. He had to overcome technical difficulties in algebraic geometry, because the Picard variety was not yet available. The fact that one can extract a relative Mordell-Weil theorem concerning only abelian varieties was given by Lang-Néron [19], whose exposition for this and the theorem of the base we have followed here. The theorem concerning sets of bounded height is also taken from [19].

As usual, our results are purely qualitative. It is of considerable interest to try to push the theory in the quantitative direction, and to give an explicit dermination for the set of generators of the Mordell-Weil group of rational points. In the geometric case, there is a result of Igusa [14] which ties it up with the second Betti number (he deals with a pencil on a surface), and with the group of principal homogeneous spaces of the abelian variety under consideration. This seem to be one of the main directions in which the theory is developing, and shows that one can expect some hard questions to arise. We may recall the following exact sequence, as in Lang-Tate [20].

Let A be an abelian variety (or even a commutative group variety) defined over the field K, and denote by A also its group of points rational over the separable closure of K. Let G be the Galois group of $\bar{K}$ over K. Then G operates on A, and if we let the cohomology be the limit cohomology from finite separable extensions, we have the exact sequence

$$0 \longrightarrow A_m \longrightarrow A \xrightarrow{m\delta} A \longrightarrow 0$$

for each integer $m > 0$ prime to the characteristic of K, giving rise to the exact cohomology sequence

$$0 \longrightarrow A_m \cap A_K \longrightarrow A_K \longrightarrow A_K \longrightarrow H^1(G, A_m) \longrightarrow H^1(G, A)$$
$$\xrightarrow{m\delta} H^1(G, A) \longrightarrow$$

the beginning which we may rewrite in the following form:

$$0 \longrightarrow A_K/mA_K \longrightarrow H^1(G, A_m) \longrightarrow H^1(G, A)_m \longrightarrow 0$$

the group on the right being the kernel in $H^1(G, A)$ of the multiplication by m. Thus $H^1(G, A)_m$ is simply the group of principal homogeneous spaces of A over K, of period m. If $A_m \subset A_K$, then we get back the Kummer theory.

Special cases when A is an elliptic curve with complex muliplication have been treated by Cassels [4]. Local results for elliptic curves have been obtained by Shafarevitch [35] and by Tate [38], who also gives a complete duality theory for abelian varieties, in preparation for global dualities. One of the main problems here is the conjecture of Shafarevitch-Tate, to the effect that those elements of $H^1(G, A)$ which split at all primes form a finite group, (conceivably modulo those which are infinitely divisible).

A propos of counting points on an abelian variety of height $\leqq B$, Néron has a very interesting conjecture (dating back to the International Congress at Edinburgh, 1958) that for a suitable choice of the projective model of an abelian variety over a number field K, having selected a basis $a_1, \cdots, a_r$ for A_K, the logarithm of the number of points of type

$$n_1a_1 + \cdots + n_ra_r$$

of height $\leqq B$ is described by a quadratic form. This would of course give much more precise information than Néron's estimate given in Theorem 5.

I can write down formally a quadratic form, of which it would be interesting to know whether it coincides with Néron's. Indeed, consider the maps

$$A \times \cdots \times A \longrightarrow A \times \cdots \times A \longrightarrow A$$

the first one being $(n_1\delta, \cdots, n_r\delta)$ and the second the sum. Using AV, Proposition 2 of Chapter IV, §1 let X be a hyperplane on A, and take its inverse image by the two maps. This leads one to consider the form

$$\sum [\eta_X(a_i + a_j) - \eta_X(a_i) - \eta_X(a_j)]n_in_j$$

where η_X is the logarithm of the height.

It would also be interesting to investigate the mapping $m\delta$ in an algebraic system of abelian varieties over, say, a number field, and to see whether the constants involved in the equivalence $H(mP) \approx H(P)^{m^2}$ can be chosen uniformly for the system.

CHAPTER VI

The Thue-Siegel-Roth Theorem

We shall give an exposition of this important theorem under axioms which are valid in number fields and function fields, namely, the product formula and a weak Riemann-Roch condition. We also need characteristic 0. In § 1 we state the theorem. In § 2 we reformulate it by a formal argument in a more manageable form (because we consider approximations at several absolute values). This should not be confused with the deep problem of simultaneous approximations at one absolute values (cf. the historical note). In § 3 we sketch the proof, basing it on two propositions. We then prove Proposition 1 in § 4 and § 5, and Proposition 2 in § 6, § 7, § 8 and § 9. Finally, in § 10, we give a geometric formulation adapted to the applications to algebraic geometry and possible generalizations to varieties of arbitrary dimensions (cf. again the historical note).

The proof of Theorem 1 is self contained and elementary (the use of the Gauss lemma of Chapter III obviously fits this description).

§ 1. *Statement of the theorem*

Throughout this chapter, we shall assume that K is a field with a proper set of absolute values M_K.

(As we deal with a fixed field K, we could omit the condition that the absolute values are well behaved: It will never be used.)

We denote by S_∞ a fixed finite, non-empty subset of M_K, containing all the archimedean absolute values, if there are any. If S is a finite subset of M_K, we denote by I_S the subset of elements $x \in K$ such that

$$|x|_v \leqq 1 \qquad \text{for all} \quad v \notin S$$

and call this the *S-integers*. If $S \supset S_\infty$ then I_S is a ring, taking into account the non-archimedean nature of the absolute values outside S. If $S = S_\infty$ we also write I_∞ or I instead of I_{S_∞} and call this simply the *integers*. We shall always assume that K is the quotient field of I_∞.

If all absolute values of M_K are non-archimedean, then the set of elements $x \in K$ such that

$$|x|_v \leqq 1 \qquad \text{for all} \quad v \in M_K$$

is a field, and all its non-zero elements have in fact absolute value 1 for all $v \in M_K$. One sees this immediately from the non-archimedean property and the product formula. This field will be called the *constant field* and will be denoted by k. Given an M_K-divisor $\mathfrak{d}$ one sees that the set $L(\mathfrak{d})$ consisting of those elements $x \in K$ satisfying $|x|_v \leqq \mathfrak{d}(v)$ is a vector space over k, again by the non-archimedean nature of the absolute values.

The following two axioms which we shall consider are very weak forms of Riemann-Roch theorems.

Axiom 1a. *There exist numbers $C_1, C_2, B_0 > 0$ depending only on K and S_∞ such that for all $B \geqq B_0$, the set of elements $x \in I_\infty$ such that*

$$|x|_v \leqq B \qquad \textit{all } v \in S_\infty$$

is finite, and the number of elements $\lambda(B)$ in the set satisfies the inequalities

$$C_1 B^N \leqq \lambda(B) \leqq C_2 B^N$$

where $N = \sum_{v \in S_\infty} N_v$.

Axiom 1b. *All absolute values of M_K are non-archimedean. There exist numbers $C_1, C_2, B_0 > 0$ and an integer $d > 0$ depending only on K and S_∞ such that for all integers $\nu > 0$ the set of elements $x \in I_\infty$ such that*

$$|x|_v \leqq B_0^\nu \qquad \textit{all } v \in S_\infty$$

is a finite dimensional vector space over the constant field k, whose dimension $l(B_0^\nu)$ satisfies the inequalities

$$C_1 \nu^d \leqq l(B_0^\nu) \leqq C_2 \nu^d$$

We note that Axiom 1a is a consequence of Theorem 2, Chapter II whenever K is a number field, and S_∞ the set of archimedean absolute values. If K is a function field in one variable over the constant field k, then Axiom 1b is a consequence of the Riemann-Roch theorem, which asserts something much more precise. In that case, if k is algebraically closed, the multiplicities turn out to be equal to 1 with suitably normalized absolute values, namely,

$$|x|_v = (1/e)^{\mathrm{ord}_v x}$$

e being the constant 2.718... .

Axiom 1b also applies to higher dimensional projective, normal varieties. One selects an absolutely irreducible hyperplane D to determine the discrete valuation at infinity. After taking a high multiple $D_0 = \nu_0 D$, one knows (from elementary facts) that $l(\nu D_0)$ is a polynomial of degree d in ν, so that Axiom 1b is satisfied. We shall use Axiom 1b to prove the next theorem. In the applications, one can also reduce the higher dimensional case to that of curves by induction (cf. [18]). In our axiomatic setup here, we may as well deal simultaneously with all cases, especially in view of the very elementary proofs which can be given for our axioms in the special cases of interest to us.

THEOREM 1. *Let K be a field of characteristic* 0, *with a proper set of absolute values M_K satisfying the product formula with multiplicities $N_v \geqq 1$. Let S_∞ be a finite, non-empty subset of M_K containing all the archimedean absolute values. Assume that K is the quotient field of I_∞ and that Axiom* 1a *or* 1b *is satisfied. Let S be a finite subset of M_K. For each $v \in S$ let α_v be algebraic over K, and assume that v is extended to $\bar{K}$ in some way. Let κ be real* > 2. *Then the elements $\beta \in K$ satisfying the approximation condition*

$$\prod_{v \in S} \inf(1, \|\alpha_v - \beta\|_v) \leqq \frac{1}{H(\beta)^\kappa}$$

have bounded height.

As in Chapter III, we use the height

$$H_K(\beta) = H(\beta) = \prod_{v \in M_K} \sup(1, \|\beta\|_v)$$

and $\|\beta\|_v = |\beta|_v^{N_v}$.

It is useful to make a few remarks concerning our statement of the theorem.

(i) The theorem is strengthened the smaller we make the left side and the larger we make the right side of our inequality. In particular, the smaller we make κ, the better the theorem. (It is known, however, that 2 is a best possible result.) The larger we make the set S, the stronger the theorem.

(ii) The right-hand side $1/H(\beta)^\kappa$ could be replaced by $C/H(\beta)^\kappa$ where C is a fixed number. Indeed, if there were β's satisfying the modified inequality with C, whose height tends to infinity, then writing $\kappa = \kappa_1 + \xi$ with $\xi > 0$ and $\kappa_1 > 2$ we have

$$\frac{C}{H(\beta)} = \frac{1}{H(\beta)^{\kappa_1}} \cdot \frac{C}{H(\beta)^{\xi}}$$

and for large $H(\beta)$, we see that $C/H(\beta)^{\xi} \leqq 1$. It is a very interesting question to estimate in number fields the number of solutions of the inequality as a function of C, asymptotically for $C \to \infty$.

(iii) Suppose that β ranges through a sequence of elements of K such that $\| \alpha_v - \beta \|_v$ tends to 0. We know from Chapter I that the extension of v to $\bar{K}$ corresponds to an embedding of $\bar{K}$ into $\bar{K}_v$. In particular, viewing α_v as embedded in $\bar{K}_v$ and thus algebraic over K_v, if α_v is approximated very closely by elements β of K, then α_v must in fact lie in K_v.

(iv) If α, α' are two distinct elements of K_v for some v, and if β approximates α, then β stays away from α'. As β approaches α, its distance from α' approaches the distance between α and α'. Hence it would add no greater generality to our statement if we took a product over several α_v for each v.

(v) There is no reason not to let β approach infinity, and we could add to the approximation condition on the left-hand side a product

$$\prod_{v \in S} \sup(1, \| 1/\beta \|_v) .$$

This version of the theorem can be reduced to the other by making a linear projective transformation $T(\beta) = (a\beta + b)/(c\beta + d)$ such that the transform of β approaches elements at a finite distance for all $v \in S$, if it converges at all. It is trivially checked that if T is a non-singular transformation with coefficients in K, then $H(T(\beta))$ is equivalent to $H(\beta)$, i.e. each is less than a constant multiple of the other.

(vi) If α is in K, and S contains one absolute value, then the theorem is trivial, and with $\kappa > 1$ instead of $\kappa > 2$. Indeed, one can even replace the finite product on the left by the complete product over all absolute values of M_K, in which case one gets the reciprocal of the height of $(\alpha - \beta)$ which is equivalent to the height of β. From this the assertion is clear.

It is known that if one puts conditions of a multiplicative nature on the β, then it is possible to lower the value of κ. This is the way one can interpret some known variants of the theorem, for instance as in [28]. Actually, these are immediate consequences of the theorem as we have stated it. As an illustration, we give the proof of a theorem of Mahler.

Corollary. *Using the same assumptions as in Theorem* 1, *and assuming in addition that the elements* β *of* K *are subject to the condition*

$$\| \beta \| = 1$$

for all $v \notin S$, *the conclusion of the theorem is valid if* $\kappa > 0$ (*instead of* $\kappa > 2$).

Proof: In view of our additional assumption, we have

$$H(\beta) = \prod_{v \in S} \sup(1, \| \beta \|_v) = \prod_{v \in S} \inf(1, \| 1/\beta \|_v)^{-1} .$$

Similarly, using $H(\beta) = H(1/\beta)$, we have

$$H(\beta) = \prod_{v \in S} \inf(1, \| \beta \|_v)^{-1} .$$

Using Remarks (iv) and (v), and assuming say, that no $\alpha_v = 0$, we know that if $\xi > 2$ then the elements β of K satisfying

$$\prod_{v \in S} \min(1, \| \alpha_v - \beta \|_v) \prod_{v \in S} \min(1, \| \beta \|_v) \prod_{v \in S} \min(1, \| 1/\beta \|_v) \leqq \frac{1}{H(\beta)^{\xi}}$$

have bounded height. Using the equalities just derived, our assertion is now obvious.

Finally, before starting the proof, it may be illuminating to consider the special case where $K = \mathbf{Q}$ and the set of absolute values is the ordinary set of absolute values. Then the multiplicities are equal to 1, and Axiom la is obviously satisfied.

When only one absolute value is considered, say the ordinary one, then the inequality has frequently been written

$$\left| \alpha - \frac{m}{n} \right| \leqq \frac{1}{|n|^{\kappa}}$$

taking m, n relatively prime integers. Since m/n comes very close to α, the two integers $| m |$, $| n |$ must have the same order of magnitude, so that instead of taking $1/| n |^{\kappa}$ for the right-hand side, one can replace $| n |$ by $\sup(| m |, | n |)$ which is precisely the height of $\beta = m/n$. This identifies the most classical form of the theorem with ours, taking into account Remark (iii), the completion being the real numbers.

Suppose we consider a p-adic absolute value. Then the term corresponding to p is sometimes written in the form

$$\min(1, |n\alpha - m|_p)$$

following Mahler. If $|\alpha - (m/n)|_p$ becomes very small, then $|m/n|_p$ cannot tend to infinity. It is bounded from above by a number depending on α. In particular, $|m|_p$ and $|n|_p$ lie between two fixed bounds depending on α, and thus having

$$\min(1, |n\alpha - m|_p)$$

is no improvement to having

$$\min\left(1, \left|\alpha - \frac{m}{n}\right|_p\right)$$

in the product.

We have thus seen how our formulation covers the classical formulations and their variants, and we are now ready to prove the theorem.

§ 2. *Reduction to simultaneous approximations*

Rather than deal with the approximation condition in the form in which we have given it, i. e. with a product, it is easier to deal with a simultaneous set of approximations for all $v \in S$. Thus we reformulate our theorem in the following manner.

THEOREM 1′. *Let the hypotheses be as in Theorem* 1. *For each* $v \in S$, *suppose a real number* $\lambda_v \geqq 0$ *is given such that*

$$\sum_{v \in S} \lambda_v = 1 \ .$$

Then the elements β *in* K *satisfying the simultaneous system of inequalities*

$$\inf(1, \|\alpha_v - \beta\|_v) \leqq \frac{1}{H(\beta)^{\kappa\lambda_v}}$$

for all $v \in S$ *have bounded height.*

It is obvious that Theorem 1 implies Theorem 1′ (take the product). We shall now show that Theorem 1′ implies Theorem 1. Suppose we have a sequence of solutions β to the approximation condition of Theorem 1, whose height tends to infinity. For such β, write

$$\inf(1, \|\alpha_v - \beta\|_v) = \frac{1}{H(\beta)^{\kappa\xi_v(\beta)}}$$

with a real number $\xi_v(\beta) \geqq 0$. Then by hypothesis,

$$\sum_{v\in S} \xi_v(\beta) \geqq 1 .$$

Now select κ' such that $\kappa > \kappa' > 2$. Choose a positive integer A such that

$$A\left(\frac{\kappa}{\kappa'} - 1\right) > s$$

where s is the number of elements of S. Using induction, and the obvious fact that $[x + y] \leqq [x] + [y] + 1$ (the bracket being the largest integer $\leqq$) we get

$$A + s \leqq A\frac{\kappa}{\kappa'} \leqq \left[\sum_{v\in S} A\frac{\kappa}{\kappa'}\xi_v(\beta)\right] \leqq \sum_{v\in S}\left[A\frac{\kappa}{\kappa'}\xi_v(\beta)\right] + s$$

whence

$$A \leqq \sum_{v\in S}\left[A\frac{\kappa}{\kappa'}\xi_v(\beta)\right].$$

Consequently there exist integers $a_v(\beta) \geqq 0$ such that

$$a_v(\beta) \leqq \left[A\frac{\kappa}{\kappa'}\xi_v(\beta)\right] \leqq A\frac{\kappa}{\kappa'}\xi_v(\beta) .$$

and $\sum a_v(\beta) = A$. From this we see that there is only a finite number of possible distributions of such integers $a_v(\beta)$, and hence, restricting our attention to a subsequence of our elements β if necessary, we can assume that the $a_v(\beta)$ are the same for all β. We write them a_v. We then put

$$\lambda_v = a_v/A$$

so that $0 \leqq \lambda_v \leqq 1$ and $\sum \lambda_v = 1$. For each β in our subsequence we have $\kappa'\lambda_v \leqq \kappa\xi_v(\beta)$, and hence these β satisfy the simultaneous system of inequalities

$$\inf(1, \|\alpha_v - \beta\|_v) \leqq \frac{1}{H(\beta)^{\kappa'\lambda_v}} .$$

We can therefore apply Theorem 1′, and have therefore shown what we wanted.

We make one more remark which is technically useful in the proof. We may assume without loss of generality that all the α_v

are integral over I_∞. Indeed, the α_v satisfy an equation

$$a_n Y^n + \cdots + a_0 = 0$$

with $a_i \in K$. (Take the product of their irreducible equations over K.) Using our assumption that K is the quotient field of I_∞ we can clear denominators, and thus assume that all a_i lie in I_∞. Multiplying this equation by a_n^{n-1}, we see that $a_n\alpha_v$ is integral over I_∞ for each v. Now if β approaches α_v, then $a_n\beta$ approaches $a_n\alpha_v$ essentially as well, and we can use a trick as in Remark (ii) to show that we may deal with $a_n\alpha_v$ instead of α_v. This proves our assertion.

§ 3. *Sketch of the proof*

Let K be a field, and $F(X) \in K[X_1, \cdots, X_m] = K[X]$ a polynomial in several variables with coefficients in K. Let $r_1, \cdots, r_m$ be integers (in $\mathbf{Z}$) $\geqq 0$ and assume that the degree of F in X_j is $\leqq r_j$. For any variable t, we have a Taylor expansion

$$F(X_1, \cdots, X_m) = \sum F^{(i)}(t, \cdots, t)(t - X_1)^{i_1} \cdots (t - X_m)^{i_m}$$

the sum being taken over $(i) = (i_1, \cdots, i_m)$, that is, m-tuples of integers satisfying $0 \leqq i_1 \leqq r_1, \cdots, 0 \leqq i_m \leqq r_m$. Each $F^{(i)}$ is a polynomial, namely

$$F^{(i)} = \frac{1}{i_1! \cdots i_m!} \partial_1^{i_1} \cdots \partial_m^{i_m} F$$

and $\partial_1, \cdots, \partial_m$ are the partial derivatives. Thus the coefficients of $F^{(i)}$ are integral multiples of the coefficients of F, since the factorials in the denominator divide the integers which will occur as a result of repeated partial differentiation. Furthermore, such integral multiples will be bounded by the product of the binomial coefficients

$$\binom{r_1}{i_1} \cdots \binom{r_m}{i_m}$$

which is itself bounded by $2^{r_1 + \cdots + r_m}$. This remark will be of use later when we estimate the absolute value of F. We also observe that the number of terms in our sum is certainly bounded by

$$(r_1 + 1) \cdots (r_m + 1)$$

which is also bounded by $2^{r_1 + \cdots + r_m}$.

We shall now state two results, to be proved in the following sections. We shall then indicate how Theorem 1′ follows from them.

If we have K, S_∞, I_∞ as described in §1, and a polynomial $G(X_1, \cdots, X_m)$ in $I_\infty[X]$ then we write

$$B_\infty(G) = \sup |G|_v$$

the sup being taken over all $v \in S_\infty$. It is a bound for the absolute values of the coefficients of G in S_∞.

We shall meet in the statements and proofs which follow certain constants depending on the original data. We number them consecutively and always indicate what they depend on.

Proposition 1. *Let K be a field with a proper set of absolute values M_K. Let S_∞ be a finite non-empty subset of M_K containing all the archimedean absolute values. Let I_∞ be its integers. Assume* Axiom 1a *or* 1b *is satisfied. Let $f(Y)$ be a polynomial in one variable in $I_\infty[Y]$ with leading coefficient* 1 *and of degree n. Let $r_1, r_2, \cdots$ be integers > 0, and let $0 < \varepsilon < 1/2$. Then for all $m > (2n/\varepsilon)^2$ (or $m > (2n/C_1\varepsilon)^2$ in case of* Axiom 1b*) there exists a polynomial*

$$G(X_1, \cdots, X_m) \neq 0$$

in $I_\infty[X]$ satisfying:

(1) $\qquad$ deg G *in* X_j *is* $\leqq r_j$.

(2) *If* $\qquad \dfrac{i_1}{r_1} + \cdots + \dfrac{i_m}{r_m} \leqq m\left(\dfrac{1}{2} - \varepsilon\right)$

then for any root α of f, we have

$$G^{(i)}(\alpha, \cdots, \alpha) = 0 \, .$$

(3) $\qquad B_\infty(G) \leqq C_3^{r_1+\cdots+r_m} \qquad$ *where* $\quad C_3 = C_3(K, S_\infty, f)$.

Remark. From now on, all constants $C_4, C_5, \cdots$ depend on K, S_∞, f only, unless otherwise specified. In each case, they can be easily evaluated.

The second proposition will give us a means of preventing a polynomial from vanishing at a point $(\beta_1, \cdots, \beta_m)$.

Proposition 2. *Let K be a field of characteristic* 0, *with a proper set of absolute values M_K satisfying the product formula with multiplicities $N_v \geqq 1$. Let m be an integer > 0. Let $0 < \delta < 1/16^{m+N}$. Let $G(X_1, \cdots, X_m) \neq 0$ be in $I_\infty[X]$ of degree $\leqq r_j$ in X_j. Assume that the r_j satisfy*

(4) $\qquad \delta r_1 > r_2, \cdots, \delta r_m > 10$ *(we say they are δ-decreasing).*

Let $\beta_1, \cdots, \beta_m$ *be elements of heights* $H_1, \cdots, H_m$ *respectively satisfying*

$$N \log 4 + 4m \leqq \delta \log H_1 \tag{5}$$

$$r_1 \log H_1 \leqq r_j \log H_j \qquad \text{for } j = 1, \cdots, m\ . \tag{6}$$

Assume also that G *satisfies*

$$B_\infty(G) \leqq H_1^{\delta r_1}\ . \tag{7}$$

Then there exists an m*-tuple* $(\nu_1, \cdots, \nu_m)$ *such that*

$$\nu_1/r_1 + \cdots + \nu_m/r_m \leqq (20)^m \delta^{(1/2)^m} \tag{8}$$

and $G^{(\nu)}(\beta_1, \cdots, \beta_m) \neq 0$.

Using Propositions 1 and 2, we show how to prove Theorem 1′.

We assume that we have elements β in K satisfying our simultaneous inequalities whose heights tend to infinity, and show that given ε, we must have $\kappa < 2 + \varepsilon'$ where ε' is small when ε is small. We can of course assume that $\varepsilon < 1/10$.

We choose $m, \delta, \beta_1, \cdots, \beta_m, r_1, \cdots, r_m$ in that order, subject to certain conditions, as follows:

We first select $m > (2n/\varepsilon)^2$ or $> (2n/C_1\varepsilon)^2$ so that we can apply Proposition 1. We then select $\delta < \varepsilon$ so that $(20)^m \delta^{(1/2)^m}$ is also less than ε.

We now select β_1 of height H_1 so large that (5) is satisfied, and that $\delta \log H_1 > \log C_3$ (the constant of (3)).

We select $\beta_2, \cdots, \beta_m$ of heights $H_2, \cdots, H_m$ respectively such that they increase logarithmically δ-rapidly, that is to say

$$2\delta \log H_2 > \log H_1$$

$$\cdots$$

$$2\delta \log H_m > \log H_{m-1}\ .$$

We choose an integer r_1 so large that

$$r_1 \delta \log H_1 > 10 \cdot \log H_j \qquad \text{for all } j = 2, \cdots, m$$

and then integers $r_2, \cdots, r_m$ such that

$$r_1 \frac{\log H_1}{\log H_j} \leqq r_j < 1 + r_1 \frac{\log H_1}{\log H_j}\ .$$

It follows immediately from the above inequalities that the r_j satisfy the condition

$$r_1 \log H_1 \leqq r_j \log H_j \leqq (1+\varepsilon) r_1 \log H_1 \tag{9}$$

even with $[1 + (1/10)\delta]$ instead of $(1 + \varepsilon)$ on the right-hand side. From this one concludes that (4) is satisfied, i.e. the r_j are δ-decreasing, namely,

$$r_j/r_{j+1} = (r_j/r_1)(r_1/r_{j+1})$$

and using (9) we get (4) at once.

We now apply Proposition 1. We had seen at the end of § 2 that we could assume the α_v to be integral, satisfying a polynomial of degree n over K with leading coefficient 1. We find a polynomial G as in that proposition. By condition (3) we conclude that

$$B_\infty(G) \leqq H_1^{\delta r_1} .$$

By Proposition 2, some derivative $G^{(\nu)} = F$ is such that $F(\beta) \neq 0$ and $\nu_1/r_1 + \cdots \nu_m/r_m \leqq \varepsilon$. Consequently, by (2) it follows that if

$$i_1/r_1 + \cdots + i_m/r_m \leqq m(1/2 - \varepsilon) - \varepsilon$$

then for any root α of f (i.e. any α_v) we have

$$F^{(i)}(\alpha, \cdots, \alpha) = 0 .$$

We estimate $\| F(\beta) \|_v$ for $v \in S$ such that $\lambda_v \neq 0$. In the expansion

$$F(\beta) = \sum F^{(i)}(\alpha_v, \cdots, \alpha_v)(\alpha_v - \beta_1)^{i_1} \cdots (\alpha_v - \beta_m)^{i_m}$$

all the terms will be 0 except those belonging to an (i) such that

$$i_1/r_1 + \cdots + i_m/r_m \geqq m(1/2 - 2\varepsilon) .$$

The total number of terms in the sum is bounded by

$$2^{r_1 + \cdots + r_m} \leqq 2^{m r_1} \leqq H_1^{\delta r_1 m}$$

since the r_j are decreasing. An upper bound for $| F^{(i)}(\alpha_v, \cdots, \alpha_v) |_v$ is certainly

$$2^{m r_1} 2^{m r_1} C_3^{m r_1} B_\infty(G) \leqq H_1^{2\delta r_1 m}$$

using the fact that we chose H_1 large. The constant C_3 need merely be a bound for the value of α_v at v. Thus putting everything together, we get

$$\| F(\beta) \|_v = H_1^{3\delta m r_1 N_v} \sup \| \alpha_v - \beta_1 \|_v^{i_1} \cdots \| \alpha_v - \beta_m \|_v^{i_m}$$

the sup being taken for those (i) satisfying the condition mentioned above. Using the hypotheses of Theorem 1′, we get

$$\| F(\beta) \|_v \leqq H_1^{3\delta m r_1 N_v} \frac{1}{(H_1^{i_1} \cdots H_m^{i_m})^{\kappa\lambda_v}} .$$

We now replace i_1 by $r_1 i_1/r_1, \cdots, i_m$ by $r_m i_m/r_m$. We then replace $H_j^{r_j}$ by $H_1^{r_1}$ using (9), making the denominator smaller and the fraction bigger. We then replace $i_1/r_1 + \cdots + i_m/r_m$ by $m(1/2 - 2\varepsilon)$ thus finally obtaining

$$\| F(\beta) \|_v \leqq H_1^{3\delta m r_1 N_v} \quad \frac{1}{H_1^{m r_1 (1/2 - 2\varepsilon) \kappa\lambda_v}} .$$

If $v \in S$ and $\lambda_v = 0$, we estimate $\| F(\beta) \|_v$ naively, and get

$$\| F(\beta) \|_v \leqq H_1^{3\delta m r_1 N_v} \sup(1, \| \beta_1 \|_v)^{r_1} \cdots \sup(1, \| \beta_m \|_v)^{r_m}$$

using the same kind of crude estimates as before for the number of terms, coefficients, etc.

If $v \notin S$, so that v is non-archimedean, then we simply get

$$\| F(\beta) \|_v \leqq \sup(1, \| \beta_1 \|_v)^{r_1} \cdots \sup(1, \| \beta_m \|_v)^{r_m} .$$

We now take the product over all $v \in M_K$. We almost get the heights of the β_j on the right, up to a finite number of factors. We make the right-hand side only bigger if we add these factors. Having done so, we use (9) which guarantees $H_j^{r_j} \leqq H_1^{r_1(1+\varepsilon)}$ and thus obtain

$$\prod_{v \in M_K} \| F(\beta) \|_v \leqq \frac{H^{6\delta m r_1 N} H_1^{m r_1 (1+\varepsilon)}}{H_1^{m r_1 (1/2 - 2\varepsilon) \kappa}} .$$

Since $F(\beta) \neq 0$, the left-hand side is equal to 1, and we take logs. Having done so, we cancel $m r_1$, and get

$$\kappa \left(\frac{1}{2} - 2\varepsilon \right) \leqq 6\delta N + 1 + \varepsilon .$$

From this we immediately get $\kappa \leqq 2 + \varepsilon'$ as desired.

§ 4. *A combinatorial lemma*

Let $r_1, \cdots, r_m$ be integers $\geqq 1$ and let $\varepsilon > 0$. The number e of sets of integers $(i_1, \cdots, i_m)$ satisfying

$$0 \leqq i_1 \leqq r_1, \cdots, 0 \leqq i_m \leqq r_m$$

and

$$\frac{i_1}{r_1} + \cdots + \frac{i_m}{r_m} \leqq m\left(\frac{1}{2} - \varepsilon\right)$$

does not exceed

$$\frac{1}{\varepsilon m^{1/2}}(r_1 + 1) \cdots (r_m + 1)$$

Proof: By induction on m. The assertion is trivial if $m = 1$. Take $m > 1$. We write

$$m\left(\frac{1}{2} - \varepsilon\right) = \frac{1}{2}(m - \lambda)$$

so that $m\varepsilon = \lambda/2$. In terms of λ our upper bound reads

$$\frac{2m^{1/2}}{\lambda}(r_1 + 1)\cdots(r_m + 1)\ .$$

Our assertion is trivial if $\lambda \leqq 2m^{1/2}$. We may therefore assume that $\lambda > 2m^{1/2}$. For each i_m, r_m fixed, we consider the solutions of

$$\begin{aligned}\frac{i_1}{r_1} + \cdots + \frac{i_{m-1}}{r_{m-1}} &\leqq \frac{1}{2}(m - \lambda) - \frac{i_m}{r_m} \\ &\leqq \frac{m-1}{2} - \frac{\lambda - 1 + 2i_m/r_m}{2}\ .\end{aligned}$$

We apply the induction hypothesis and sum over i_m. For this, we first make a computation:

$$\begin{aligned}\sum_{i=0}^{r} \frac{2}{\lambda - 1 + 2i/r} &= \sum_{i=0}^{r}\left[\frac{1}{\lambda - 1 + 2i/r} + \frac{1}{\lambda + 1 - 2i/r}\right] \\ &= \sum_{i=0}^{r} \frac{2\lambda}{\lambda^2 - (1 - 2i/r)^2} \\ &\leqq \sum_{i=0}^{r} \frac{2\lambda}{\lambda^2 - 1} = (r + 1)\frac{2\lambda}{\lambda^2 - 1}\end{aligned}$$

We now let $i = i_m$ and $r = r_m$. Our number e is therefore bounded by

$$\begin{aligned}&\sum \frac{2(m-1)^{1/2}}{\lambda - 1 - 2i_m/r_m}(r_1 + 1)\cdots(r_{m-1} + 1) \\ &\quad\leqq \frac{2\lambda(m-1)^{1/2}}{\lambda^2 - 1}(r_1 + 1)\cdots(r_m + 1)\ .\end{aligned}$$

We must therefore check whether

$$\frac{2\lambda(m-1)^{1/2}}{\lambda^2-1} \leqq \frac{2m^{1/2}}{\lambda}\,,$$

or equivalently, whether

$$\lfloor(m-1)/m]^{1/2} \leqq (\lambda^2-1)/\lambda^2\,.$$

Observe that $(1-1/m)^{1/2} \leqq 1-1/2m$ (square the right-hand side). It suffices therefore that $1-1/2m \leqq 1-1/\lambda^2$. But this is true in view of our original hypothesis on λ, and our lemma is proved.

§ 5. *Proof of Proposition* 1

We have to use either Axiom 1a or 1b. Let us start with 1a, and indicate afterwards the modifications of language necessary to use Axiom 1b instead.

The number of polynomials $F(X_1, \cdots, X_m)$ such that

(10) $$\deg F \text{ in } X_j \text{ is } \leqq r_j$$

(11) $$F \in I_\infty[X] \quad \text{and} \quad B_\infty(F) \leqq B$$

is at least $(C_1B)^{N(r_1+1)\cdots(r_m+1)}$. For each set of integers (i) satisfying the condition of (2), we let $RF^{(i)}$ be the remainder of the division of $F^{(i)}(Y, \cdots, Y)$ by $f(Y)$. The degree of $F^{(i)}(Y, \cdots, Y)$ (which is a polynomial in one variable) is at most $r_1 + \cdots + r_m$. Using the remarks at the beginning of § 3, and long division, we find by an inductive estimate that

(12) $$B_\infty(RF^{(i)}) \leqq C_5^{r_1+\cdots+r_m} B_\infty(F)$$

where C_5 depends on f and S_∞ only. In this way, each polynomial satisfying (10) and (11) gives rise to a set of e remainders

$$(RF^{(i)_1}, \cdots, RF^{(i)_e})$$

where e is bounded as in the preceding section. Each remainder is of degree $< n$.

By Axiom 1a, the number of such sets of remainders is at most

(13) $$[C_5^{r_1+\cdots+r_m}B]^{Nne}\,.$$

We want two distinct polynomials with the same set of remainder, say F, G. Then $F - G$ will have remainders 0 for all (i) satisfying (2), and thus the conclusion of (2) will be satisfied. Furthermore, $B_\infty(F-G)$ will be at most $2B$, and thus (3) will also be satisfied. Condition (1) is trivially satisfied since we started with (10).

We shall certainly get our two distinct polynomials if

$$(C_1B^N)^{(r_1+1)\cdots+(r_m+1)} > [C_5^{r_1+\cdots+r_m}B]^{Nne}\,, \tag{14}$$

and this will certainly be true if we replace e by the upper estimate derived in the preceding section. The product $(r_1+1)\cdots(r_m+1)$ in the exponent of both sides can then be cancelled. Let $C_6 = C_5^N$. It will suffice to get

$$C_1B^N > [C_6^{r_1+\cdots+r_m}B^N]^{n/\varepsilon m^{1/2}}\,. \tag{15}$$

If now $m > (2n/\varepsilon)^2$, then it will suffice to get

$$C_1{}^NB > [C_6^{r_1+\cdots+r_m}B^N]^{1/2}\,, \tag{16}$$

and thus it will in fact suffice to get

$$B^{N/2} > C_7^{r_1+\cdots+r_m}\,, \tag{17}$$

with a suitable C_7. Letting $C_8 = C_7^{2/N}$ we get a condition of the type desired to satisfy (3). This concludes the proof, using Axiom 1a.

To use Axiom 1b, we have to count linear dimensions.

The polynomials $F(X)$ satisfying (10) and

$$F \in I_\infty[X] \qquad \text{and } B_\infty(F) \leqq B_0^\nu \tag{11b}$$

are a vector space over the constants k of dimension at least equal to $(r_1+1)\cdots(r_m+1)C_1\nu^d$.

For each set of integers (i) satisfying the condition of (2), we find the same estimate (12) (without bothering to take into account the fact that we are dealing with non-archimedean absolute values). The map

$$F \to (RF^{(i)_1}, \cdots, RF^{(i)_e})$$

s now a k-linear map. By Axiom 1b, the dimension over k of the vector space which is the image of this map is at most

13b) $$ne[\nu_0(r_1+\cdots+r_m)+\nu]^d\,.$$

he integer ν_0 is determined so that $C_5 \leqq B_0^{\nu_0}$. We now want the ernel of our k-linear map to be larger than 0, and for this it will suffice to have the inequality of dimensions:

$$(r_1+1)\cdots(r_m+1)C_1\nu^d > n\frac{1}{\varepsilon\nu_m}(r_1+1)\cdots(r_m+1) \cdot [\nu_0(r_1+\cdots r_m)+\nu]^d\,. \tag{14b}$$

We cancel the product of the $(r_j + 1)$, and take $m > (2n/C_1\varepsilon)^2$. We then see that it suffices to achieve

$$\nu^d > \frac{1}{2}[\nu_0(r_1 + \cdots + r_m) + \nu]^d$$

and taking a d-th root, it is enough to get

$$\nu\left(1 - \frac{1}{2^{1/d}}\right) > \nu_0(r_1 + \cdots + r_m) .$$

Putting this condition in terms of $B_\infty(F)$, we see that our proposition is proved.

§6. *Wronskians*

Let R be a ring. A *derivation* d of R is a linear map of R into itself satisfying

$$d(xy) = x \cdot dy + y \cdot dx .$$

Those elements x of R such that $dx = 0$ form a subring (because $d(1) = 0$) containing the prime ring, and are called the ring of *constants*. This ring of constants is a field if R is a field. A *differential ring* is a ring together with a finite set of derivations $d_1, \cdots, d_n$ which commute with each other (as operators on R). Its subring of constants is by definition the intersection of the rings of constants of each d_j.

Example. Let k be a field, and R the polynomial ring $k[X_1, \cdots, X_n]$. Let d_j be the partial derivative with respect to X_j, taken formally. Then R is a differential ring, and if k has characteristic 0, it is the field of constants.

If R is a differential ring, and $d_1, \cdots, d_n$ its derivations, we can form monomials

$$d_1^{i_1} \cdots d_n^{i_n}$$

where i_ν are integers $\geqq 0$. These monomials can be viewed as linear operators on R in the obvious way. The *degree* of the above monomial is by definition $i_1 + \cdots + i_n$. We denote by $D(r)$ the set of monomials of degree $< r$, where r is an integer $\geqq 1$.

The following proposition will be applied to the case of polynomials in several variables, as above. I am indebted to Kolchin for the formulation and proof in its present degree of generality.

Proposition 3. *Let F be a differential field, k the field of constants* *for the derivations* $d_1, \cdots, d_n$ *of* F. *Let* $D(r)$ *be the monomials of degree* $< r$ *in the derivations. Let* $x_1, \cdots, x_n$ *be elements of* F. *If they are linearly dependent over* k, *then*

$$\det(\theta_i x_j) = 0$$

for all operators $\theta_1, \cdots, \theta_n \in D(r)$ *and any* r. *Conversely, if this determinant is* 0 *for all choices of* $\theta_i \in D(i)$, *then they are linearly dependent over* k.

Proof: The first assertion is trivial. We prove the second by induction on n, and the assertion is trivial for $n = 1$. We assume $n > 1$. Observe that $\theta_1 = 1$. We consider determinants:

$$\begin{vmatrix} \theta_1 x_1 & \theta_1 x_2 \cdots \theta_1 x_n \\ & \cdots \\ \theta_n x_1 & \theta_n x_2 \cdots \theta_n x_n \end{vmatrix} \qquad \theta_i \in D(i)\,.$$

If all the subdeterminants of size $n - 1$ above the last line are 0 for all choices of $\theta_i \in D(i)$, then any $n - 1$ among the x's are linearly independent by induction, and we are done. Thus we may assume that some determinant of size $n - 1$ above the last line is not 0, say

$$\begin{vmatrix} x_1 & x_2 & \cdots & x_{n-1} \\ \theta'_2 x_1 & \theta'_2 x_2 & \cdots & \theta'_2 x_{n-1} \\ & \cdots & & \\ \theta'_{n-1} x_1 & \cdots & & \theta'_{n-1} x_{n-1} \end{vmatrix} \neq 0$$

for some choice of $\theta'_i \in D(i), i = 1, \cdots, n - 1$.

We complete this determinant to one with rows having n elements. Then these rows are linearly independent over F since the smaller rows are already linearly independent over F. Hence any row

$$(\theta x_1, \cdots, \theta x_n) \qquad \theta \in D(n)$$

is linearly dependent over F on our expanded $n - r$ rows, and thus the matrix

$$\begin{bmatrix} x_1 & \cdots & x_n \\ \theta'_2 x_1 & \cdots & \theta'_2 x_n \\ & \cdots & \\ \theta'_{n-1} x_{n-1} & \cdots & \theta'_{n-1} x_n \\ \theta x_1 & \cdots & \theta x_n \\ \varphi x_1 & \cdots & \varphi x_n \\ \cdots & \cdots & \cdots \end{bmatrix}$$

has rank $n-1$ for any choice of $\theta, \varphi, \cdots$ in $D(n)$. It therefore has column rank equal to $n-1$, and hence there exist elements $a_1, \cdots, a_n \in F$ with $a_n = 1$ such that for all $\theta \in D(n)$ we have

$$a_1\theta x_1 + \cdots + a_n\theta x_n = 0 \tag{18}$$

because the last column can be expressed as a linear combination of the first $(n-1)$ columns. Let d be any one of the derivations d_i. Applying d to (18) we get

$$a_1(d\theta x_1) + \cdots + a_n(d\theta x_n) + (da_1)\theta x_1 + \cdots + (da_n)\theta x_n = 0\ .$$

In this last expression, let $A(\theta)$ and $B(\theta)$ denote the first sum involving the $d\theta$, and the second sum respectively. We have $A(\theta) + B(\theta) = 0$. If θ happens to lie in $D(n-1)$ then $d\theta$ lies in $D(n)$ and hence $A(\theta) = 0$ by (18). This holds in particular for everyone of the derivations θ_j' and gives $B(\theta_j') = 0$ for all j. This is a set of $n-1$ equations involving $n-1$ variables

$$da_1, \cdots, da_{n-1}$$

because $a_n = 1$ and $da_n = 0$. The determinant of the coefficients is not zero, and hence $da_1 = \cdots = da_{n-1} = 0$. Since d was any one of the derivations d_i it follows that $a_1, \cdots, a_{n-1}$ are constants. Letting $\theta = 1$ in (18) proves what we wanted.

§ 7. *Factorization of a polynomial*

Although it is not true that a polynomial in several variables factorizes into a product of polynomials in fewer variables, we shall see, nevertheless, that a suitable derivative of it does. This will allow us to carry out an induction argument later.

Let K be a field of characteristic 0. We consider the polynomial ring $K[X_1, \cdots, X_m] = K[X]$ with the standard partial derivatives. By a *differential operator* D we mean

$$D = \frac{1}{i_1! \cdots i_m!}\partial_1^{i_1} \cdots \partial_m^{i_m}$$

and we call $i_1 + \cdots + i_m$ the *degree* of the operator. Applied to any polynomial G, it yields another polynomial whose coefficients are integral multiples of those of G.

LEMMA 1. *Let K be a field of characteristic* 0. *Let* $G(X) =$

$G(X_1, \cdots, X_m)$ be a non-zero polynomial in $K[X]$ in m variables, $m \geqq 2$, of degree $\leqq r_j$ in X_j for each $j = 1, \cdots, m$. Then there exists an integer l satisfying

$$1 \leqq l \leqq r_m + 1 \tag{19}$$

and differential operators $D, \cdots, D_{l-1}$, on the variables $X_1, \cdots, X_{m-1}$, of degrees $\leqq 0, \cdots, l-1$, respectively, such that if

$$F(X_1, \cdots, X_m) = \det\left(D_\nu \frac{1}{\mu!} \partial_m^\mu G\right) \qquad (\nu, \mu = 0, \cdots, l-1)$$

then $F \neq 0$, and

$$F(X_1, \cdots, X_m) = U(X_1, \cdots, X_{m-1}) V(X_m)$$

where U, V are polynomials, U is of degree at most lr_j in X_j for $j = 1, \cdots, m-1$ and V is of degree at most lr_m in X_m.

Proof: We consider all representations of G in the form

$$G(X) = \varphi_0(X_m)\psi_0(X_1, \cdots, X_{m-1}) + \cdots + \varphi_{l-1}(X_m)\psi_{l-1}(X_1, \cdots, X_{m-1})$$

where the φ_ν and ψ_μ are polynomials with coefficients in K, of degree at most r_j in X_j. Such a representation is possible, for instance, with $l-1 = r_m$ and $\varphi_\nu(X_m) = X_m^\nu$. From all such representations, we select the one for which l is least. Then

$$\varphi_0(X_m), \cdots, \varphi_{l-1}(X_m)$$

are linearly independent over K, for, if not, one would immediately deduce another representation of G as above with a smaller l. Similarly,

$$\psi_0(X_1, \cdots, X_{m-1}), \cdots, \psi_{l-1}(X_1, \cdots, X_{m-1})$$

are linearly independent over K. Also $1 \leqq l \leqq r_m + 1$.

We now apply Proposition 3 to these two sets of polynomials. The θ's in Proposition 3 were taken with coefficient 1. Since we are in characteristic 0, putting a factorial in the denominator results in no loss of generality, and hence there exist differential operators $D_0, \cdots, D_{l-1}$, in the variables $X_1, \cdots, X_{m-1}$, of degrees $\leqq 0, \cdots, l-1$, respectively, such that

$$U(X_1, \cdots, X_{m-1}) = \det(D_\mu \psi_\rho) \qquad (\mu, \rho = 0, \cdots, l-1)$$

is not 0. Similarly,

$$V(X_m) = \det\left(\frac{1}{\nu!} \partial_m^\nu \varphi_\rho\right) \qquad (\nu, \rho = 0, \cdots, l-1)$$

is not 0. Multiplying these two determinants, we see that $UV = F$. The degrees of U, V in the X_j obviously satisfy the required condition, since the determinants are $l \times l$.

We now estimate the coefficients of F in terms of those of G.

Lemma 2. *Let the notation be as in the previous lemma. Let v be an absolute value on K. If v is non-archimedean then $|F|_v \leqq |G|_v^l$. If v induces the ordinary absolute value on* $\mathbf{Q}$, *then*

$$|F|_v \leqq [(r_1+1)\cdots(r_m+1)]^l l!\, 2^{l(r_1+\cdots+r_m)} |G|_v^l .$$

Proof. If v is non-archimedean, our assertion is obvious. If v induces the ordinary absolute value on $\mathbf{Q}$, we regard G as a sum of $(r_1+1)\cdots(r_m+1)$ terms, each of the form

$$a_{(i)} X_1^{i_1}\cdots X_m^{i_m}$$

where $|a_{(i)}| \leqq |G|$. We can develop our determinant into a sum of $[(r_1+1)\cdots(r_m+1)]^l$ determinants, the general element in each such determinant being of the form

$$a_{(i)} D_\mu \frac{1}{\nu!} \partial_m^\nu X_1^{i_1}\cdots X_m^{i_m} .$$

By the remarks at the begining of § 3, we see that the effect of the differential operators on the estimates is at most to multiply them by $2^{r_1+\cdots+r_m}$. Hence the coefficients of each of the $l!$ terms in the expansion of an individual determinant have absolute values not exceeding

$$[|G|\, 2^{r_1+\cdots+r_m}]^l$$

and our assertion follows.

§ 8. *The index*

Let K be a field, $F(X_1, \cdots, X_m)$ a polynomial in $K[X]$, and $F \neq 0$. Let $r_1, \cdots, r_m$ be numbers > 0, and $\beta_1, \cdots, \beta_m$ elements of K. To simplify the notation, we denote by $r_{(m)}$ and $\beta_{(m)}$ the vectors of the r's and β's respectively.

Let t be a new variable. Consider the expansion

$$F(\beta_1 + X_1 t^{1/r_1}, \cdots, \beta_m + X_m t^{1/r_m})$$
$$= \sum F^{(i)}(\beta_1, \cdots, \beta_m) X_1^{i_1}\cdots X_m^{i_m} t^{i_1/r_1+\cdots+i_m/r_m} .$$

We can view the right-hand side as a polynomial in fractional powers of t. We let

$$\theta = \operatorname{ind}_{r_{(m)},\, \beta_{(m)}} F$$

be the smallest number occurring as an exponent of t on the right-hand side such that t^{θ} has a non-zero coefficient, and call it the *index* of F at $\beta_{(m)}$ relative to $r_{(m)}$. Then θ is also the minimum of the numbers $(i_1/r_1 + \cdots + i_m/r_m)$ such that $F^{(i)}(\beta_1, \cdots, \beta_m) \neq 0$.

We note that the index is always $\geqq 0$, and equals 0 if and only if $F(\beta_1, \cdots, \beta_m) = 0$.

It follows immediately from the definitions that the index satisfies the following properties:

IND. 1. *If F, G are two polynomials, neither of which is* 0, *then*

$$\operatorname{ind}(F+G) \geqq \min(\operatorname{ind} F, \operatorname{ind} G)$$
$$\operatorname{ind}(FG) = \operatorname{ind} F + \operatorname{ind} G\ .$$

(because the degree of a product is equal to the sum of the degrees). *If F is a polynomial in $X_1, \cdots, X_{m-1}$ and G is a polynomial in X_m, then*

$$\operatorname{ind}_{r_{(m)},\, \beta_{(m)}}(FG) = \operatorname{ind}_{r_{(m-1)} \beta_{(m-1)}}(F) + \operatorname{ind}_{r_m, \beta_m}(G)\ .$$

IND 2. *The index of*

$$\partial^{\nu_1} \cdots \partial_m^{\nu_m} F$$

at $\beta_{(m)}$ relative to $r_{(m)}$ is at least equal to

$$\operatorname{ind}_{r_{(m)}, \beta_{(m)}}(F) - \nu_1/r_1 - \cdots - \nu_m/r_m\ ,$$

for any integers $\nu_1, \cdots, \nu_m \geqq 0$, provided the derived polynomial is not identically 0.

IND 3. *Let l be a number > 0, and $lr_{(m)} = (lr_1, \cdots, lr_m)$. Then*

$$\operatorname{ind}_{lr_{(m)}, \beta_{(m)}}(F) = \frac{1}{l} \operatorname{ind}_{r_{(m)}, \beta_{(m)}}(F)\ .$$

We now wish to compare the indices of G and F in Lemma 1.

LEMMA 3. *Let the notation be as in Lemma* 1. *Let $0 < \delta \leqq 1$. Assume that*

$$\delta r_1 \geqq r_2, \cdots, \delta r_m \geqq 10\ . \tag{20}$$

Let $\beta_1, \cdots, \beta_m$ be elements of K, and put $\operatorname{ind} = \operatorname{ind}_{r_{(m)}, \beta_{(m)}}$. *Let* $\theta = \operatorname{ind}(G)$. *Then*

$$\theta^2 \quad \textit{or} \quad \theta \leqq \frac{4}{l} \operatorname{ind}(F) + 4\delta\ .$$

Proof: Consider a differential operator of the form

$$D_\mu = \frac{1}{i_1! \cdots i_{m-1}!}\partial_1^{i_1}\cdots\partial_{m-1}^{i_{m-1}}$$

on $X_1, \cdots, X_{m-1}$ of degree $\leqq l-1$. If the polynomial

$$D_\mu \frac{1}{\nu!}\partial_m^\nu R$$

does not vanish identically, its index at $\beta_{(m)}$ relative to $r_{(m)}$ is at least

$$\theta - i_1/r_1 - \cdots - i_{m-1}/r_{m-1} - \nu/r_m \, .$$

We make this number smaller if we replace the r_j $(j = 1, \cdots, m-1)$ by r_{m-1}. Hence this index is at least

$$\theta - \frac{l-1}{r_m - 1} - \frac{\nu}{r_m} \geqq \theta - \frac{r_m}{r_m - 1} - \nu/r_m \, .$$

Since $r_m/r_{m-1} \leqq \delta$, and since the index is never negative, it must be at least

$$\max(0, \theta - \nu/r_m) - \delta \, .$$

If we expand the determinant giving the expression for F, we obtain for F a sum of $l!$ terms, the typical term being of the form

$$\pm (D_{\mu_0}G)\left(D_{\mu_1}\frac{1}{1!}\partial_m G\right)\cdots\left(D_{\mu_{l-1}}\frac{1}{(l-1)!}\partial_m^{l-1}G\right)$$

where the D_μ have degree $\leqq l-1$. By one of the basic properties of the index, if such a term does not vanish identically its index is at least

$$\sigma = \sum_{\nu=0}^{l-1}\max(0, \theta - \nu/r_m) - l\delta \, .$$

Since F is a sum of such terms, and is not 0, it follows that $\operatorname{ind}(F) \geqq \sigma$. We must now solve for θ in terms of $\operatorname{ind}(F)$. We can assume $\theta r_m > 10$, for if not, we have

$$\theta \leqq 10/r_m \leqq \delta$$

and our inequality is certainly satisfied.

If $\theta r_m < l$ then we have

$$\begin{aligned}\sum_{\nu=0}^{l-0}\max(0, \theta - \nu/r_m) &= 1/r_m \sum_{\nu=0}^{l-1}\max(0, \theta r_m - \nu)\\ &\geqq 1/r_m \sum_{0\leqq\nu\leqq[\theta r_m]}([\theta r_m] - \nu)\\ &\geqq \frac{[\theta r_m]([\theta r_m]+1)}{2r_m}\\ &\geqq \frac{[\theta r_m]^2}{2r_m}\,.\end{aligned}$$

If a is a number > 10, then $(a-1)^2/2 > a^2/3$. Consequently, since $\theta r_m > 10$, we have by a trivial computation

$$[\theta r_m]^2/2r_m \geqq (\theta r_m)^2/3r_m = \theta^2 r_m/3\,.$$

Thus in this case, $\operatorname{ind}(F) \geqq \theta^2 r_m/3 - l\delta$ and

$$\theta^2 \leqq \frac{l}{r_m}\frac{3}{l}\operatorname{ind}(F) + \frac{3l}{r_m}\delta \leqq \frac{4}{l}\operatorname{ind}(F) + 4\delta\,,$$

since $r_m \geqq 10$ and $l \leqq r_m + 1$.

On the other hand, if $\theta r_m \geqq l$ we have

$$\sum_{\nu=0}^{l=1}\max(0, \theta - \nu/r_m) = \sum_{\nu=0}^{l-1}(\theta - \nu/r_m) \geqq l\theta/2\,.$$

From this our assertion is again obvious, and this concludes the proof.

§ 9. *Proof of Proposition* 2

We come to the proof of Proposition 2. We shall in fact prove a slightly stronger version of it. If K is a field with a proper set of absolute values M_K satisfying the product formula with multiplicities $N_v \geqq 1$, and $F(X_1, \cdots, X_m) \in K[X]$ is a polynomial, then we define as usual

$$H(F) = \prod_{v\in M_K} |F|_v^{N_v} = \prod_{v\in M_K} \|F\|_v\,.$$

PROPOSITION 4. *Let K be a field of characteristic* 0, *with a proper set of absolute values M_K satisfying the product formula with multiplicities $N_v \geqq 1$. Let S be a finite subset of M_K containing the archimedean absolute values, and let $N = \sum_{v\in S} N_v$. Let*

$$0 < \delta < 1/16^{m+N}\,. \tag{21}$$

Let $r_j\,(j = 1, \cdots, m)$ be integers satisfying

$$\delta r_1 > r_2, \cdots, \delta r_m > 10 . \tag{4}$$

Let $G(X_1, \cdots, X_m)$ *be a non-zero polynomial in* $K[X]$ *of degree* $\leqq r_j$ *in* X_j *and let* $\beta_j (j = 1, \cdots, m)$ *be elements of* K *with* $H(\beta_j) = H_j$. *Assume that:*

$$N \log 4 + 4m \leqq \delta \log H_1 \tag{22}$$

$$r_1 \log H_1 \leqq r_j \log H_j \tag{23}$$

$$H(G) \leqq H_1^{N\delta r_1} . \tag{24}$$

Then

$$\operatorname{ind}_{r_{(m)}, \beta_{(m)}}(G) \leqq (20)^m \delta^{(1/2)^m} .$$

Remark. What we want, of course, is that if δ is small, the coefficients of G are small, and the r_j are δ-decreasing, then the index of G is small. The hypotheses are made precise in order to load the induction.

Proof: We prove Proposition 4 by induction on m. Suppose first $m = 1$. Write $G(X) = (X - \beta)^e F(X)$ and suppose $\deg G \leqq r$. By Gauss' lemma we have

$$H[(X - \beta)^e] H(F) \leqq H(G) 4^{rN} .$$

For each absolute value in M_K, it is clear that

$$|(X - \beta)|^e \leqq |(X - \beta)^e|$$

and hence $H(\beta)^e \leqq H[(X - \beta)^e]$. From this and the fact $H(F) \geqq 1$ we get $H(\beta)^e \leqq H(G) 4^{rN}$, whence

$$e \log H(\beta) \leqq \log H(G) + rN \log 4 .$$

By hypothesis and conditions (24) and (22) we get

$$e \leqq N\delta r + rd$$

and consequently the index $\operatorname{ind}_{r,\beta}(G) = e/r$ satisfies

$$e/r \leqq \delta(N + 1) \leqq 6\delta^{1/2} ,$$

if we use (21). Thus the case $m = 1$ is proved.

Let $m > 1$. We use Lemma 1 and get a polynomial $F = UV$ as in this lemma. The estimate for its coefficients obtained in Lemma 2 can now be transformed as follows, using (4):

$$|F|_v \leqq 2^{mr_1 l} 2^{r_1 l} 2^{mr_1 l} |G|_v^l$$

replacing $(r_j + 1)$ by 2^{r_j} and $l!$ by $l^l \leqq 2^{r_1 l}$. Using (22) and (24) this yields

$$H(F) \leqq 8^{Nmr_1 l} H(G) \leqq H_1^{2\delta N r_1 l} .$$

We also know that the degree of F in X_j is $\leqq lr_j$.

Since $F = UV$ and the variables are separated, one can verify immediately that $H(F) = H(U)H(V)$, and hence we get the same estimate for $H(U)$ and $H(V)$ that we have just obtained for $H(F)$, because the height is always $\geqq 1$. Both U and V have degrees $\leqq r_j$ in X_j. In (24) we can replace G by U or V and δ by 2δ, so that it remains valid (using (23) for V). In (4) we can also replace δ by 2δ and r_j by lr_j so that it remains valid. Thus the induction hypothesis is satisfied for each one of U, V with lr_j and 2δ.

Using Property 1 of the index, we get by induction

$$\operatorname{ind}_{r_{(m-1)}, \beta_{(m-1)}}(U) \leqq (20)^{m-1}(2\delta)^{(1/2)^{m-1}}$$
$$\operatorname{ind}_{r_m, \beta_m}(V) \leqq (20)(2\delta)^{(1/2)} .$$

The index for V is certainly at most equal to the index for U. The index of F at $\beta_{(m)}$ at $\beta_{(m)}$ relative to $lr_{(m)}$ is the sum of these two. Using Lemma 3 and the fact that $4\delta \leqq 4(2\delta)^{1/2}$ we get

$$\theta^2 \qquad \text{or} \qquad \theta \leqq (4 \cdot 3 \cdot \sqrt{2})(20)^{m-1}(2\delta)^{(1/2)^{m-1}}$$

from which we get

$$\theta \leqq (20)^m (2\delta)^{(1/2)^m}$$

thereby proving our proposition, and Roth's theorem.

§ 10. *A geometric formulation of Roth's theorem*

Throughout this section, K is a field satisfying the hypotheses of Theorem 1, so that this theorem is valid for it.

We reformulate Theorem 1 slightly, as an intermediate step to the final statement of Theorem 2 below.

Let $G(Y)$ be a polynomial in $K[Y]$ (one variable) and assume that the multiplicity of its roots is at most r for some integer $r > 0$. Let S be a finite subset of M_K. Let $C > 0$ be a number, and $\kappa > 2$. Then the solutions β in K of the inequality

$$\prod_{v \in S} \inf(1, \| G(\beta) \|_v) \leqq \frac{C}{H(\beta)^{\kappa r}}$$

have bounded height.

To prove this, we may assume that G has leading coefficient 1, and say

$$G(Y) = \prod_{i=1}^{d} (Y - \alpha_i)^{e_i}$$

is a factorization in $\bar{K}$. We extend our absolute values of S to $\bar{K}$ in any way. The expression on the left-hand side of our inequality is greater or equal to

$$\prod_{v \in S} \prod_{i=1}^{d} \inf(1, \| \beta - \alpha_i \|_v)^{e_i}$$

which is itself greater or equal to the expression obtained by replacing e_i by r for each i. Now we are in the situation of Theorem 1, taking into account the remark (iv) following it, namely, the solutions β of the inequality

$$\prod_{v \in S} \prod_{i=1}^{d} \inf(1, \| \beta - \alpha_i \|_v)^{r} \leqq \frac{1}{H(\beta)^{\kappa r}}$$

have bounded height, hence so do the solutions of our original inequality.

We observe that Theorem 1 is in fact equivalent to this theorem, for if α is algebraic over K, we take for $G(Y)$ its irreducible equation over K. Its roots occur with multiplicity 1, and if β approximates α, then it stays away from the other roots, at a distance greater than some fixed constant, once it is sufficiently close to α. We can do the same for each α_v, and thus we see that the implication is trivial.

THEOREM 2. *Let W be a complete non-singular curve defined over K. Let z and y be two non-constant functions in $K(W)$, and let r be the largest of the multiplicities of the zeros of z. Assume that y is defined at all the zeros and poles of z and gives an injective mapping of this set of zeros into $\bar{K}$. Let κ be a number > 2 and let $C > 0$. Then the points $Q \in W_K$ which are not poles or zeros of z and are such that*

$$\prod_{v \in S} \inf(1, \| z(Q) \|_v) \leq \frac{C}{H(y(Q))^{\kappa r}}$$

have bounded height H_y.

Proof: Without loss of generality, we may assume that $K(W) =$

$K(z, y)$ (otherwise, replace W by a non-singular model of $K(z, y)$). We may also assume that the absolute values

$$|\, y(Q) \,|_v$$

are bounded for all $v \in S$. Indeed, if there is an infinite sequence of points Q whose height $H_y(Q) = H(y(Q))$ tends to infinity and which satisfies the above inequality, but with $|\, y(Q) \,|_v$ unbounded for some $v \in S$, we then consider a non-singular linear transformation on y, say $y' = (ay + b)/(cy + d)$, with coefficients in K, which has the same property as y in the statement of the theorem, but such that $|\, y'(Q) \,|_v$ is bounded for all $v \in S$, and a subsequence of points Q. $H_{y'}$ is then equivalent to H_y (trivial by direct verification, or use Property 3 of heights). If S has only one absolute value v, then we can actually take $1/y$ instead of y.

Let Φ be the set of zeros of z. Since y has no pole in Φ, it is integral over the local ring $\mathfrak{o}$ of the point $z = 0$ in the function field $K(z)$. Let $F(Y)$ be its irreducible equation over $\mathfrak{o}$. Then

$$F(Y) \equiv G(Y) \qquad (\text{mod } z)$$

where $G(Y)$ is a polynomial with coefficients in K with leading coefficient 1 and mod z means modulo the maximal ideal of $\mathfrak{o}$ generated by z.

By hypothesis, y induces an injection $Q \to y(Q)$ of Q into $\bar{K}$. The multiplicity of a root of $G(Y)$ is thus $\leqq$ the multiplicity of a point on W in the inverse image of the point $z = 0$, since it is the multiplicity of a zero of z. (One can see this formally for instance as follows: Let $(y^{(1)})$ be an affine generic point of w over K all of whose coordinates are integral over $\mathfrak{o}$. If $(y^{(1)}, \cdots, y^{(m)})$ is a complete set of conjugates of $(y^{(1)})$ over $K(z)$, then the cycle on W which is the inverse image of $z = 0$ consists of a specialization $(\bar{y}^{(1)}, \cdots, \bar{y}^{(m)})$ of $(y^{(1)}, \cdots, y^{(m)})$ over $z \to 0$. The conjugates of y correspond to the conjugates $(y^{(1)}, \cdots, y^{(m)})$, and one can then use IAG, Theorem 2 of Chapter I, § 4 applied to the polynomial $F(Y)$.

We can write

$$F(Y) = G(Y) + zA(z, Y)$$

where $A(z, Y)$ is a polynomial in Y with coefficients in $\mathfrak{o}$. Since $A(0, Y)$ is defined, so is $A(z(Q), Y)$ for small values of $\|\, z(Q) \,\|_v$, which is all that we need to consider. Thus the values

$$\|\, A(z(Q), y(Q)) \,\|_v$$

remain bounded since we may assume that $\|y(Q)\|_v$ remains bounded for all $v \in S$. Since $F(y) = 0$, we get an estimate for $G(y(Q))$, namely,

$$\|G(y(Q))\|_v \leqq C' \|z(Q)\|_v$$

for some constant C' and all $v \in S$. Taking the usual product, we get

$$\begin{aligned} \prod_{v \in S} \inf(1, \|G(y(Q))\|_v) &\leqq C'' \inf(1, \|z(Q)\|_v) \\ &\leqq \frac{C''}{H(y(Q))^{\kappa r}} \end{aligned}$$

which leaves us in the situation discussed at the beginning of this section, and concludes the proof.

Actually, it will be even more convenient to make use of the projection property of Chapter IV, § 4. Of course, in proving these properties of heights, we worked with a ground field F and absolute values which were well behaved in order to state our theorems for the absolute heights. Here, it does not matter since we shall work over a fixed field K (we incorporate the multiplicities N_v directly into the data). Properties of heights clearly remain valid under our present set up. Thus from the projection property, we get:

THEOREM 3. *Let W be a projective non-singular curve defined over K. Let ψ be a non-constant function in $K(W)$ and let r be the largest of the multiplicities of the poles of ψ. Let κ be a number > 2, C a number > 0. Let S be a finite subset of M_K. Then the points $Q \in W_K$ which are not among the poles of ψ, and are such that*

$$\prod_{v \in S} \sup(1, \|\psi(Q)\|_v) \geqq C \cdot H(Q)^{\kappa r}$$

have bounded height.

Proof: We let Φ be the set of zeros and poles of ψ. We make a projection to get a function y so that we can apply the preceding theorem. We have also found it convenient to invert the formula, taking $z = 1/\psi$.

In the applications, we could either apply the formula directly, or work with only one absolute value in S. For simplicity, we shall follow this second alternative in the next chapter.

Historical Note:

We refer to Schneider's book for a complete historical note on the Liouville-Thue-Siegel-Mahler-Dyson-Schneider-Roth theorem. Mahler's contribution was to have included several primes (i.e. taking the product).

The proof follows that given by Roth without any essential change. A deeper understanding of the situation would come from three direction:

(1) Making the result quantitative. For instance, letting $\kappa = 2 + \xi$ and finding a bound for the height of β in terms of ξ as $\xi \to 0$ (cf. [8]). Another possibility: Allowing $C/H(\beta)^\kappa$ on the right-hand side of the approximation condition and finding asymptotic bounds for $H(\beta)$ in terms of $C \to \infty$.

(2) Dealing with simultaneous approximations. Here, it seems to be generally conjectured that the following statement should be true. Let $\alpha_1, \cdots, \alpha_n$ be algebraic over K, and assume that $1, \alpha_1, \cdots, \alpha_n$ are linearly independent over K. Then the height of points $(1, \beta_1, \cdots, \beta_n) = \beta$ with $\beta_i \in K$ such that the simultaneous conditions

$$| \alpha_i - \beta_i | \leqq \frac{1}{H(\beta)^\kappa}$$

hold is bounded if $\kappa > 1 + 1/n$. (Over the rationals, one puts all the β_i over a common denominator, with relatively prime numerators, and one can therefore replace $H(\beta)$ by this denominator.)

(3) Generalizing the geometric formulation to, say, projective non-singular varieties V. The condition $\kappa > 2$ should be replaced by a similar one (instead of 2, perhaps $1 + n$ where $n = \dim V$), and the multiplicity r by a multiplicity which must be a biholomorphic invariant, so that it remains the same in unramified coverings. This is essential if one is to apply the theorem to isogenies of abelian varieties, (cf. the comments at the end of Chapter VII). In this connection the case of linear functions has been investigated, and and we refer the reader to a paper of Schmidt for asymptotic results concerning approximations to almost all numbers. (W. Schmidt, *A metrical theorem in diophantine approximation*, Canadian Journal of Mathematics Vol. 12 (1960 pp. 619–631, Theorem 2.)

CHAPTER VII

Siegel's Theorem

If C is an affine curve defined over a ring R finitely generated over $\mathbf{Z}$, and if its genus is $\geqq 1$, then C has only a finite number of points in R. This is the central result of the chapter. We shall also give a relative formulation of it for a curve defined over a ring which is a finitely generated algebra over an arbitrary field k of characteristic 0. In that case, the presence of infinitely many points in R implies that the curve actually comes from a curve defined over the constant field and that its points are of a special nature (excluding possibly a finite number).

In § 1, we put together Roth's theorem and the weak Mordell-Weil theorem to show that integral points are of bounded height. It is actually more convenient (although not really more general) to deal with a non-constant function φ whose values $\varphi(P)$ lie in a ring of the above type.

In § 2, we then consider the two standard examples and prove the statements of absolute and relative finiteness.

To deal with curves of genus 0, it will be convenient to use the decomposition theorem (although in this special case, it would need less machinery to prove it). Suppose we deal with a projective non-singular variety defined over a field K with a proper set of absolute values which we assume to be discrete for a moment. For instance, we may take a number field together with the set of all its proper non-archimedean absolute values, or a subset of it, excluding only a finite number of them. Then, given an element of K, we have its M_K-divisor or ideal (the same thing in this case). Given a projective non-singular variety V defined over K, let $\varphi \in K(V)$ and suppose that we write the divisor of φ in terms of its decomposition into prime rational divisors of V over K:

$$(\varphi) = \sum m_\lambda T_\lambda$$

where the T_λ are distinct and the m_λ are integers. If P is any point of V_K which does not lie in the support of the zeros or poles of φ, then $\varphi(P)$ is an element of K. We shall describe its ideal decomposition in terms of the T_λ. In fact, with each prime divisor

T of V over K and each point P, we shall associate an integral M_K-ideal $\mathfrak{a}_T(P)$ such that

$$(\varphi(P)) = \sum m_\lambda \mathfrak{a}_{T_\lambda}(P)$$

up to a fixed number of components depending only on φ and not on P. Furthermore, if V is of dimension 1, then the different $\mathfrak{a}_{T_\lambda}$ are "relatively prime" up to a fixed ideal.

In developing the theory, we do not restrict ourselves to discrete absolute values, and the hypothesis that V is projective non-singular is used only to guarantee the existence of the theory. Except for that, we can proceed axiomatically and assume only that V is normal and complete.

We shall apply the decomposition to curves of genus 0 to classify those which have infinitely many integral points. If φ is a non-constant function on V, V is assumed to be of dimension 1 as above, and if R is a subring of K all of whose elements are integral for the valuations in M_K, then $\varphi(P) \in R$ means that for each pole T of φ, our ideal $\mathfrak{a}_T(P)$ lies between two fixed bounds. This will imply that the function φ is of a very particular type, which will be determined in § 4.

§ 1. *Height of integral points*

Throughout this section, we shall deal with a field $\mathbf{F}$ *with a proper set of absolute values* $M_{\mathbf{F}}$ *satisfying the product formula. Let* K *be a finite extension of* $\mathbf{F}$, *with its set of absolute values* M_K. Let $v \in M_K$. We shall say that K *satisfies Roth's theorem for* v if the following assertion is true:

Given a projective non-singular curve U *defined over* K, *let* $\psi \in K(U)$ *be a non-constant function, and let* r *be the largest of the orders of the poles of* ψ. *Let* κ *be a number* >2, *and* $c > 0$. *Then the points* $Q \in U_K$ *which are not poles of* ψ, *and are such that*

$$|\psi(Q)|_v \geqq cH(Q)^{\kappa r}$$

have bounded height.

Here, and in the rest of this section, we have put $H = H_{M_K}$.

Using the weak Mordell-Weil theorem, our purpose is to improve this statement for curves of genus $\geqq 1$. We shall prove:

THEOREM 1. *Let* K *be as above, satisfying Roth's theorem for one of the absolute values* $v \in M_K$. *Let* C *be a complete non-singular curve* *of genus* $\geqq 1$ *defined over* K, *let* J *be its Jacobian, also defined over* K, *let* m *be an integer* >0, *prime to the characteristic such that* J_K/mJ_K *is finite. Let* φ *be a non-constant function in* $K(C)$, *and let* ρ, c *be numbers* >0. *Then the height of points* P *in* C_K *which are not poles of* φ *and such that*

$$| \varphi(P) |_v \geqq cH(P)^\rho$$

is bounded.

In view of Property 4 of heights, we observe that if a set of points has bounded height in one projective embedding, then it has bounded height in every projective embedding. Therefore we need not specify the embedding in statements such as the above.

If C is a complete non-singular curve defined over K, of genus $\geqq 1$, and with a rational point, then its Jacobian J is also defined over K, and we may assume that C is contained in it under a canonical mapping. We shall take J with a fixed projective embedding, and we shall take on C the induced embedding. The height $H = H_K$ is taken relative to this embedding.

Proposition 1. *Let* m *be an integer* >0, *unequal to the characteristic of* K, *and assume that* J_K/mJ_K *is finite. Let* $\mathfrak{S}$ *be an infinite set of rational points of* C *in* K. *Then there exists an unramified covering* $\omega : U \to C$ *defined over* K, *an infinite set of rational points* $\mathfrak{S}'$ *of* U *in* K, *such that* ω *induces an injection of* $\mathfrak{S}'$ *into* $\mathfrak{S}$, *and a projective embedding of* U *over* K *such that* $H \circ \omega$ *is quasi-equivalent to* H^{m^2}. (*Of course, in* $H \circ \omega$ *the* H *refers to the height on* C, *while in* H^{m^2} *it refers to the height on* U.)

Proof: Let $a_1, \cdots, a_n$ be representatives of cosets of J_K/mJ_K. Infinitely many $P \in \mathfrak{S}$ lie in the same coset, and so there exists one point, say a_1, and infinitely many points $Q \in J_K$ such that $mQ + a_1$ lies in $\mathfrak{S}$. We let $\mathfrak{S}'$ be this infinite set of points Q. The covering $\omega : J \to J$ given by $\omega u = mu + a_1$ is unramified, and its restriction U to C also gives an unramified covering of C, which is an irreducible covering of the same degree as ω. The inverse image of a point in C lies in U. Thus $\mathfrak{S}'$ is actually a subset of U_K. Restricting one's attention to a subset of $\mathfrak{S}'$ guarantees that ω induces an injection on this subset.

To prove the relation concerning the height, we may work on the Jacobian itself since we take U in the projective embedding induced by that of J. If X is a hyperplane section of J, then

$$\omega^{-1}(X) = (m\delta)^{-1}X_{-a_1} \equiv m^2X$$

where $\equiv$ is algebraic equivalence. But $H_{\omega^{-1}(X)} \sim H_X \circ \omega$ by property 2 of heights (which is trivial), and $H_{\omega^{-1}(X)}$ is quasi-equivalent to $H_{m^2X} \sim H_K^{m^2}$ by Property 4. Our proposition is therefore proved.

Although it is not difficult to prove that the restriction of our covering ω to C is of the same degree as ω (cf. [15] or [33]) it is actually irrelevant here, since we could just as well use one of the components of this restriction that contains infinitely many points of S'. Such a component must then be defined over K.

We note that $\varphi(P) = \varphi(\omega(Q))$. Let ψ be the function $\varphi \circ \omega$ on U. Let κ be a number >2, and let m be large enough so that $m^2\rho > \kappa r$. Then for a suitable constant $c > 0$, our inequality becomes

$$|\psi(Q)|_v \geqq cH(Q)^{\kappa r}$$

and we can apply Roth's theorem, because the orders of the poles of ψ on U are the same as those of φ on C, since our covering is unramified. This concludes the proof of Theorem 1.

THEOREM 2. *Let K, C, φ be as in Theorem 1. Let S be a finite, non-empty subset of M_K, containing all the archimedean primes and assume that K satisfies Roth's theorem for each $v \in S$. Let R be a subring of K all of whose elements are integral for each valuation $v \notin S$, and let $\mathfrak{R}$ be the subset of C_K consisting of those points P which are not poles of φ, and such that $\varphi(P) \in R$. Then the points of $\mathfrak{R}$ have bounded height.*

Proof: Assume the contrary, and let $\mathfrak{R}_1$ be a subsequence of points of $\mathfrak{R}$ whose height tends to infinity. By assumption, we have $|\xi|_v \leqq 1$ for $v \notin S$ and $\xi \in R$. Hence for all $P \in \mathfrak{R}_1$ we get

$$H(\varphi(P)) = \prod_{v \in S} \sup(1, |\varphi(P)|_v)^{N_v}.$$

Let $N = [K : \mathbf{F}]$ and let s be the number of elements of S. Since $N_v \leqq N$, we have at most Ns terms in our product, of type

$$\sup(1, |\varphi(P)|_v) .$$

Consequently, for each $P \in \mathfrak{R}_1$ there exists one v in S such that

$$|\varphi(P)|_v \geqq H(\varphi(P))^{1/Ns} .$$

Hence there exists an infinite subset $\mathfrak{R}_2$ of $\mathfrak{R}_1$ such that for some $v \in S$ and all points $P \in \mathfrak{R}_2$ we have

$$|\varphi(P)|_v \geqq H(\varphi(P))^{1/N\varepsilon}.$$

In view of the corollary to Property 4 of heights, we can compare $H(\varphi(P))$ and $H(P))$. If d is the degree of C in its given projective embedding, and φ is viewed as a map into $\mathbf{P}^1$, then there is a number $c_1 > 0$ depending on ε such that for P in C_K we have

$$H(P)^{r/d-\varepsilon} \leqq c_1 H(\varphi(P)) .$$

Combining this with our previous inequality, we see that there is a number $\rho > 0$ such that for some $v \in S$ and all $P \in \mathfrak{R}_2$ we have for suitable $c_2 > 0$:

$$|\varphi(P)|_v \geqq c_2 H(P)^{\rho}$$

and we can apply Theorem 1 to get our contradiction.

The special cases of interest of Theorem 2 are number fields and function fields. In the case of number fields, we take for R a finitely generated ring over $\mathbf{Z}$, and then we conclude that the set of points P such that $\varphi(P)$ lies in R, is actually finite. This will be generalized to rings of finite type over $\mathbf{Z}$, including transcendental elements in the next section.

In the case of function fields, we shall consider a function field K over a constant field k of characteristic 0. The set of absolute values is that of Example 2, Chapter II, § 1. All the conditions of Theorem 2 are satisfied, except that J_K/mJ_K need not be finite. However, in view of the fact that the height is geometric, i.e. does not depend on the constant field (Chapter III, Proposition 6, § 3), we may go over to the function field $K\bar{k}$ over $\bar{k}$, i.e. we may without loss of generality assume that k is algebraically closed. In that case, we can apply Theorem 2 of Chapter VI, taking into account that B_k (and hence τB_k) is divisible, and thus that J_K/mJ_K is indeed finite.

In the next section, we shall give a more precise description of our sets of bounded height and sets of integral points.

§ 2. *Finiteness theorems*

If K is a function field over a constant field k we shall pick a model as in Example 2 of Chapter II to compute heights. If C is a complete non-singular curve defined over K, and a subset of points of C_K has bounded height in some projective embedding, then we recall that it has bounded height in every projective embedding (say by Property 4 of heights).

PROPOSITION 2. *Let K be a function field over a constant field k of characteristic* 0. *Let C be a complete non-singular curve of genus $\geqq 1$ defined over K. Let $\mathfrak{R}$ be an infinite subset of C_K of bounded height. Regard C as embedded in its Jacobian J over K, and let (B, τ) be a K/k-trace of J. Then τ is an isomorphism. The points of $\mathfrak{R}$ lie in a finite number of cosets of τB_k. If infinitely many of them lie in one coset, and so are of type $a + \tau b$ where a is some point of J_K and b ranges over an infinite subset of B_k, then $C_0 = \tau^{-1}(C_{-a})$ is defined over k, and τ induces an isomorphism of C_0 onto C_{-a}.*

Proof: We may assume that J and B are embedded in projective space. We see from the Corollary of Theorem 6 of Chapter VI, § 5 that the points of $\mathfrak{R}$ lie in a finite number of cosets of τB_k. Assume that infinitely many lie in the coset $a + \tau B_k$. We know that τ establishes an isomorphism between B and τB by AV, Corollary 2 of Theorem 9, Chapter VIII. Since infinitely many points of τB_k lie in C_{-a}, it follows that their K-closure in τB or J is precisely C_{-a}. Put $C_0 = \tau^{-1}(C_{-a})$. Then C_0 is a curve contained in B, and τ induces an isomorphism τ_0 of C_0 onto C_{-a}. Furthermore, C_0 contains infinitely many points b of B rational over k. It is then a trivial matter to conclude that C_0 is defined over k, because these infinitely many points are both k- and K-dense in C_0. Since τB contains a translation of C, it follows that $\tau B = J$ is the Jacobian, i.e. τ is an isomorphism.

Remark. If we do not assume characteristic 0 in the preceding proposition, then τ is merely bijective, and C_0 may be defined over a purely inseparable extension of k.

COROLLARY. *Let K, k be as above. Let C be a complete non-singular curve of genus $\geqq 2$ defined over K, and let $\mathfrak{R}$ be an infinite subset of C_K consisting of points of bounded height. Then there exists a curve C_0 defined over k and a birational transformation $T: C_0 \rightarrow C$ defined over K such that all but a finite number of points of $\mathfrak{R}$ are images under T of rational points of C_0 in k.*

Proof: Since the genus is $\geqq 2$, the curve cannot be equal to any translation of itself in its Jacobian. (If it were, so would the divisor Θ, and it isn't, even up to linear equivalence, by AV Theorem 3 of Chapter VI, § 3.) Hence there can only be one coset having infinitely many points of C, and we apply the proposition. Combining our corollary with the results obtained in the previous section, we get:

Theorem 3. *Let K be a function field over a constant field k of characteristic* 0. *Let R be a subring of K of finite type over k. Let C be a complete non-singular curve of genus $\geqq 1$ defined over K, and φ a non-constant function on C, also defined over K. Let $\mathfrak{R}$ be the subset of C_K consisting of those points P such that $\varphi(P) \in R$. If $\mathfrak{R}$ is infinite, then there exists a curve C_0 defined over k and a birational transformation $T : C_0 \to C$ defined over K. If the genus is $\geqq 2$, then all but a finite number of points of $\mathfrak{R}$ are images under T of points of C_{0k}. If the genus is* 1, *then the points of $\mathfrak{R}$ lie in a finite number of cosets of $T(C_{0k})$.*

Our next theorem gives the absolute result.

Theorem 4. *Let K be a field of finite type over* **Q**, *and R a subring of K of finite type over* **Z**. *Let C be a projective non-singular curve of genus $\geqq 1$ defined over K, and let φ be a non-constant function in $K(C)$. Then there is only a finite number of points $P \in C_K$ which are not poles of φ and such that $\varphi(P) \in R$.*

Proof: If R is algebraic over **Z**, this is simply Theorem 2 of § 1 together with the finiteness of points of bounded height in a given number field. We shall reduce the general case to this one by a specialization argument, made possible by the previous theorem.

Suppose that R is infinite. Let k be the algebraic closure of **Q** in K. Then by Theorem 3, C is birationally equivalent over K to a curve C_0 defined over k. If the genus of C is 1, we restrict our attention to an infinite subset of points of R which lie in the same coset of $T(C_{0k})$. Then without loss of generality, we may assume $C = C_0$, and that we have infinitely many points of C in K such that $\varphi(P) \in R$, where φ is a function defined over K. We shall now prove our theorem by induction on the dimension of K over **Q**.

Let F be a subfield of K containing k such that the dimension of K over F is 1. There exists a discrete valuation ring $\mathfrak{o}$ of K containing F and R whose residue class field $E = \mathfrak{o}/\mathfrak{m}$ is finite over F and such that the reduction φ' of φ mod $\mathfrak{m}$ is a non-constant function $\varphi' : C \to \mathbf{P}^1$ (of the same degree as φ). For any point Q of C_K, we get a specialized point Q' in C_E and $\varphi'(Q') = \varphi(Q)'$ using the compatibility of intersections and reductions, i.e. formally, using the graphs:

$$[\Gamma_\varphi \cdot (Q \times \mathbf{P}^1)]' = \Gamma'_\varphi \cdot (Q' \times \mathbf{P}^1)$$

the left-hand side being $Q \times \varphi(Q)$ and the right-hand side being $Q' \times \varphi'(Q')$. This yields infinitely many points Q' of C_E such that $\varphi'(Q')$ lies in the ring R', the image of R in the homomorphism $\mathfrak{o} \to \mathfrak{o}/\mathfrak{m}$. Since E is of finite type over $\mathbf{Q}$, of dimension one less than that of K, and since R' is still of finite type over $\mathbf{Z}$, this concludes the proof.

§ 3. *Theorem of decomposition*

Throughout this section, K is a field with a proper set of absolute values M_K. V is a complete, normal variety defined over K. All points, functions, and divisors on V will be assumed to be rational over $\bar{K}$. All places to be considered will be assumed to be $\bar{K}$-valued places of $\bar{K}(V)$ over $\bar{K}$. We use v to denote absolute values in M_K and w to denote absolute values in $M_{\bar{K}}$.

If φ is a function on V (and thus in $\bar{K}(V)$ according to our conventions), and π a place, then either $\pi(\varphi)$ is ∞ in which case we put

$$| \pi(\varphi) |_w = \infty \qquad w \in M_{\bar{K}}$$

or $\pi(\varphi)$ is finite and $| \pi(\varphi) |_w$ has the usual meaning. We agree that ∞ is greater than any real number, and that as usual, $1/0 = \infty$ while $1/\infty = 0$.

If P is a point which does not lie in the divisor of poles of φ, then for any place π_P which induces P on V, $\pi_P(\varphi)$ is the same and is equal to $\varphi(P)$. If P does not lie in the divisor of zeros of φ, then $\pi_P(\varphi) = \infty$ if P lies in the divisor of poles, and is finite $\neq 0$ if P lies neither in the divisor of zeros or poles of φ, i.e. if $P \notin \operatorname{supp}(\varphi)$. These criteria will be applied constantly in the sequel. (Cf. IAG, Proposition 3 of Chapter VI, § 1.)

By a *distribution* we shall mean a function $\mathfrak{d}(\pi, w)$ of places π and absolute values w which can be expressed in the form

$$\mathfrak{d}(\pi, w) = \sup_i \inf_j | \pi(\varphi_{ij}) |_w$$

where $\varphi_{ij}(i = 1, \cdots, r$ and $j = 1, \cdots, s)$ are functions on V. We then write $\mathfrak{d} = \mathfrak{d}_\varphi = \mathfrak{d}_{\varphi_{ij}}$ and we say that the distribution is attached to the set φ_{ij}, or φ.

It is clear that if $\mathfrak{d}(\pi, w)$ is a distribution attached to the set φ_{ij}, then $\mathfrak{d}(\pi, w)^{-1}$ is also a distribution, attached to the set $\psi_{ji} = 1/\varphi_{ij}$.

If $\mathfrak{d}(\pi, w)$ and $\mathfrak{d}'(\pi, w)$ are two distributions, attached to the sets φ_{ij} and $\psi_{\nu\mu}$ respectively, and if their product is defined (i.e. in no case is one of the factors 0 while the other is ∞) then the product $\mathfrak{d}(\pi, w)\mathfrak{d}'(\pi, w)$ is also a distribution, attached to the set $\omega_{i\nu, j\mu} = \varphi_{ij}\psi_{\nu\mu}$, i.e. the product is equal to

$$\sup_{i,\nu} \inf_{j,\mu} | \pi\varphi_{ij}\psi_{\nu\mu}) |_w \, .$$

This assertion is an immediate consequence of the definitions.

A distribution $\mathfrak{d}$ is said to be *rational* over a field E (algebraic over K) if $\mathfrak{d} = \mathfrak{d}_{\varphi_{ij}}$ for some set of functions φ_{ij} rational over E.

Let $\mathfrak{d}$ be a distribution and P a point of V. We shall say that $\mathfrak{d}$ is *defined* at P if it satisfies the following conditions:

(i) $\mathfrak{d}(\pi, w)$ is the same for all places π inducing P on V, and is then denoted by $\mathfrak{d}(P, w)$.

(ii) $\mathfrak{d}(\pi, w) \neq \infty$ for any w, and if $\mathfrak{d}$ is rational over a finite extension E of K, and P is also rational over E, then $\mathfrak{d}(\pi, w)$ depends only on the absolute value induced by w on E, and as such is an M_E-divisor.

The next proposition gives an important criterion for a distribution to be defined at a point.

Proposition 3. *Let T be a divisor on V and assume that we can write*

$$(\varphi_{ij}) = X_i - Y_j + T$$

where X_i, Y_j are divisors $\geqq 0$, and the Y_j have no point in common. Then the distribution $\mathfrak{d}_\varphi$ is defined at any point $P \notin supp(T)$, and either P is a common point of all the X_i, in which case $\mathfrak{d}_\psi(P, w) = 0$ for all w, or the X_i have no point in common, and then

$$\mathfrak{d}_\varphi(P, w) = sup_{i'}\, inf_{j'} \mid \varphi_{i'j'}(P) \mid_w \neq 0, \infty$$

where i' ranges over the indices such that $P \notin supp(X_{i'})$ and j' ranges over the indices such that $P \notin supp(Y_{j'})$.

Proof: Assume $P \notin \text{supp}(T)$. Given an index i, we have two cases:

Case 1. $P \notin \text{supp}(X_i)$. Then for some index j' we know that $P \notin \text{supp}(Y_{j'})$ and hence $\varphi_{ij'}(P)$ is defined. Thus

$$\inf_j \mid \pi_P(\varphi_{ij}) \mid_w = \inf_{j'} \mid \varphi_{ij'}(P) \mid_w \, .$$

Case 2. $P \in \text{supp}(X_i)$. Then $\inf_j \mid \pi_P(\varphi_{ij}) \mid_w = 0$. Taking the sup over i, our assertions are now clear.

Corollary. *Let the notation be as in the previous proposition and* *assume in addition that T is positive. If $P \in supp(T)$, then $\mathfrak{d}_\varphi$ is defined at P and equal to* 0.

We recall that if c is an M_K-divisor, we put $c(w) = c(v)$ for any $w \in M_{\bar{K}}$ extending some $v \in M_K$.

Let $\mathfrak{d}$ and $\mathfrak{d}'$ be two distributions. They will be said to be *equivalent* (and we write $\mathfrak{d} \sim \mathfrak{d}'$) on a set of points U if they are both defined on U, and if there exist two M_K-divisors c_1 and c_2 such that

$$c_1(w)\mathfrak{d}(P, w) \leqq \mathfrak{d}'(P, w) \leqq c_2(w)\mathfrak{d}(P, w)$$

for all $P \in U$ and all $w \in M_{\bar{K}}$. We shall give a criterion for two distributions to be equivalent. For this purpose, we need a lemma.

LEMMA 1. *Let T be a divisor on V, $T \geqq 0$, and let φ_{ij} be functions such that*

$$(\varphi_{ij}) = X_i - Y_j + T$$

where the X_i, Y_j are positive divisors, the X_i have no point in common, and neither do the Y_j. Then there exists an M_K-divisor c such that for all points P we have for all $w \in M_{\bar{K}}$

$$\mathfrak{d}_\varphi(P, w) \leqq c(w).$$

Proof: It will suffice to prove that for each i,

$$\inf_j |\pi(\varphi_{ij})|_w \leqq c(w) ,$$

and this relation is equivalent to

$$\sup_j |\pi(1/\varphi_{ij})|_w \geqq c(w)^{-1} .$$

On the other hand, we have for each i,

$$(1/\varphi_{ij}) = Y_j - X_i - T .$$

Since $T \geqq 0$, our functions $1/\varphi_{ij}$ have the property of Lemma 1, Chapter IV, § 2 and we are finished.

COROLLARY. *If $T = 0$ in the lemma, then $\mathfrak{d}_\varphi$ is equivalent to* 1.

Let T be a divisor on V, and assume that we can find functions φ_{ij} and positive divisors X_i, Y_j such that the X_i have no point in common, the Y_j have no point in common, and

$$(\varphi_{ij}) = X_i - Y_j + T .$$

We shall then say that we can *attach a distribution to T*. If $\psi_{\nu\mu}$, X'_ν, Y'_μ is another set of such functions and divisors, then $\mathfrak{d}_\varphi$ and

$\mathfrak{d}_\psi$ are equivalent on $V - \text{supp}(T)$. Indeed, each one is $\neq 0, \neq \infty$ and $\mathfrak{d}_\varphi/\mathfrak{d}_\psi$ is the distribution $\mathfrak{d}_\omega$ where $\omega_{i\mu,j\nu} = \varphi_{ij}/\psi_{\mu\nu}$. Thus

$$(\omega_{i\mu,j\nu}) = X_i + Y'_\mu - Y_j - Y'_\nu$$

(we may obviously assume that none of the φ_{ij} or $\psi_{\nu\mu}$ is the function 0). We can therefore apply the corollary of Lemma 1 to conclude that $\mathfrak{d}_\omega$ is equivalent to 1, and hence that $\mathfrak{d}_\varphi$ and $\mathfrak{d}_\psi$ are equivalent.

We are interested in distributions only up to equivalence. If we can attach a distribution to a divisor T, we shall denote by $\mathfrak{d}_T$ any distribution in the equivalence class of distributions $\mathfrak{d}_\psi$ obtained in the above manner.

We shall now list systematically the properties of distributions associated with divisors.

Property 1. *If $T \geqq 0$ and we can attach a distribution to T, then we can take $\mathfrak{d}_T \leqq 1$, and $\mathfrak{d}_T$ is defined and >0 on the complement of supp(T). It is defined and equal to 0 on supp(T).*

This follows immediately from Lemma 2, from which we see that $\inf(\mathfrak{d}_\psi, 1)$ is equivalent to $\mathfrak{d}_\psi$.

Property 2. *If T and T' are two divisors to which we can attach distributions $\mathfrak{d}_T$ and $\mathfrak{d}_{T'}$ respectively, then we can attach a distribution to $T + T'$ and*

$$\mathfrak{d}_{T+T'} \sim \mathfrak{d}_T \mathfrak{d}_{T'}$$

on the complement of supp$(T) \cup$ supp(T').

This assertion is obvious, simply by taking the products of two sets of functions φ_{ij} and $\psi_{\nu\mu}$ giving rise to the distributions $\mathfrak{d}_T$ and $\mathfrak{d}_{T'}$ respectively.

Property 3. *Let f be a function on V. Let U be the open set where f is defined. Then we have a distribution $\mathfrak{d}_f$ defined on U, and such that for each $P \in U$, we have*

$$\mathfrak{d}_f(P, w) = |f(P)|_w .$$

This is trivial, putting all $\varphi_{ij} = f$.

Property 4. *If T, T' are two divisors to which we can attach distributions $\mathfrak{d}_T$ and $\mathfrak{d}_{T'}$ respectively, and if they are linearly equivalent, i.e. if there is a function f such that*

$$(f) = T' - T$$

then $\mathfrak{d}_f$ *is equivalent to* $\mathfrak{d}_{T'}/\mathfrak{d}_T$ *on the complement of* $supp(T) \cup supp(T')$.

Indeed, if φ_{ij} are used to define $\mathfrak{d}_T$ as above, then we can clearly use the functions $f\varphi_{ij}$ to define $\mathfrak{d}_{T'}$.

Property 4 is known as the *decomposition theorem.*

PROPERTY 5. *If* $T_1, \cdots, T_m$ *are positive divisors on* V *to which we can attach distributions, then* $sup_\lambda \mathfrak{d}_{T_\lambda}$ *is a distribution. If the* T_λ *have no point in common, then* $sup_\lambda \mathfrak{d}_{T_\lambda}$ *is equivalent to* 1.

Suppose that we have for each index λ,

$$(\varphi_{i_\lambda, j_\lambda}) = X_{i_\lambda} - Y_{j_\lambda} + T_\lambda$$

as usual. Let the pairs (λ, i_λ) form a new set of indices ν, and the pairs (λ, j_λ) form a new set of indices μ. Then

$$\begin{aligned} \sup_\lambda \mathfrak{d}_{T_\lambda}(\pi, w) &= \sup_\lambda \sup_{i_\lambda} \inf_{j_\lambda} |\,\pi(\varphi_{i_\lambda, j_\lambda})\,|_w \\ &\geqq \sup_\lambda \sup_{i_\lambda} \inf_\lambda \inf_{j_\lambda} |\,\pi(\psi_{\nu\mu})\,|_w \\ &= \sup_\nu \inf_\mu |\,\pi(\psi_{\nu\mu})\,|_w \end{aligned}$$

where $\psi_{\nu\mu} = \varphi_{i_\lambda, j_\lambda}$. This shows that our $\sup_\lambda \mathfrak{d}_{T_\lambda}$ is a distribution. If the T_λ have no point in common, then by the Corollary of Lemma 1, this distribution is equivalent to 1. Indeed,

$$(\psi_{\nu\mu}) = X'_\nu - Y'_\mu$$

where $X'_\nu = X_{i_\lambda} + T_\lambda$ and $Y'_\mu = Y_{j_\lambda}$ so our assumptions imply that the corollary applies.

Finally, we shall give a criterion for the possibility of attaching distributions to divisors. If T is a divisor on V, and K_T its smallest field of rationality containing K, we shall say that we can *attach a distribution to* T *rationally* over K_T if the functions φ_{ij} and the divisors X_i, Y_j of Proposition 3 can be selected rational over K_T.

THEOREM 5. *Let* V *be a projective non-singular variety defined over the field* K *having a proper set of absolute values* M_K. *Then for every divisor* T *on* V *rational over* $\bar{K}$, *one can attach a distribution rationally over* K_T.

Proof: We just use Lemma 2 of Chapter IV, § 3 so that $T + X \sim Y$ where X, Y are both ample. We can choose them both rational over K_T. Then we simply pick hyperplanes $X_i \sim X$ and $X_j \sim Y$ rational over K_T and having the required properties (i.e. the X_i without points in common, and the Y_j also). (Of course the X_i and Y_j are hyperplanes in different projective embeddings.)

As mentioned previously, the main application of the theorem of decomposition will be to curves. If we translate Properties 4 and 5 to a context where all the absolute values of M_K are discrete, we get:

THEOREM 6. *Let K be a field with a proper set M_K of absolute values, all of them discrete. Let V be a complete non-singular curve defined over K. To each prime rational divisor T of V over K, and to each point $P \in V_K$, one can associate an integral M_K-ideal $\mathfrak{a}_T(P)$ having the following property: If φ is a function in $K(V)$, and*

$$(\varphi) = \sum m_\lambda T_\lambda$$

is its divisor where the T_λ are distinct prime divisors of V rational over K, then there exist two M_K-ideals $\mathfrak{b}_1$ and $\mathfrak{b}_2$, depending only on φ, such that for any $P \in V_K$ which is not a zero or pole of φ, we have

$$\mathfrak{b}_1 + \sum m_\lambda \mathfrak{a}_{T_\lambda}(P) \leqq (\varphi(P)) \leqq \mathfrak{b}_2 + \sum m_\lambda \mathfrak{a}_{T_\lambda}(P) .$$

If T and T' are distinct prime divisors, then there exists an integral M_K-ideal $\mathfrak{b}$ such that for all $P \in V_K$ we have

$$0 \leqq inf\, [\mathfrak{a}_T(P), \mathfrak{a}_{T'}(P)] \leqq \mathfrak{b} .$$

It is convenient to reformulate our notion of equivalence in the special case of Theorem 6. Therefore let K and V be as in this theorem. Suppose that we have a mapping $\mathfrak{a}$ from V_K to the M_K-ideals, so for each $P \in V_K$, $\mathfrak{a}(P)$ is an M_K-ideal. Two such maps $\mathfrak{a}, \mathfrak{a}'$ are called *equivalent* if there exist two M_K-ideals $\mathfrak{b}_1$, $\mathfrak{b}_2$ such that for all $P \in V_K$ we have

$$\mathfrak{b}_1 + \mathfrak{a}(P) \leqq \mathfrak{a}'(P) \leqq \mathfrak{b}_2 + \mathfrak{a}(P) .$$

In Theorem 6, we see that our maps $\mathfrak{a}_T$ preserve the usual operations of divisors and ideals, i.e. the sum, sup, and inf, up to equivalence.

We observe finally that the theory of distributions can be developed without a “constant field” as the beginnings of it are done in [40]. One could also formulate the decomposition theorem in that case, i.e. essentially for a variety scheme all of whose fibers are also varieties and for absolute values on the function fields of these fibers. We leave this as an exercise to the reader.

§ 4. *Curves of genus* 0

As a typical example of the arguments which follow, let us *consider a curve*

$$ax + by = 1$$

with coefficients $a, b \neq 0$ *in a field* K *of characteristic* 0. *Let* Γ *be a finitely generated multiplicative subgroup of* K^*. *Then there is only a finite number of solutions of the above equation with* $x, y \in \Gamma$.

Proof: Let n be an integer $\geqq 3$. If there were infinitely many, then there would be infinitely many in one coset of Γ/Γ^n, which we could write in the form

$$x = a_1\xi^n, \qquad \text{and} \qquad y = b_1\eta^n \ .$$

This would give infinitely many solutions in $\Gamma \times \Gamma$ of the equation

$$aa_1\xi^n + bb_1\eta^n = 1$$

which has genus $\geqq 1$. We can view these solutions as lying in the ring finitely generated by Γ over $\mathbf{Z}$, thereby contradicting Theorem 4 of § 2.

By a *linear torus* (*torus* for short) we shall mean a group variety which is isomorphic to a product of multiplicative groups. We note that given a finitely generated subgroup of a commutative group variety, there always exists a field of finite type over the prime field over which all the points of this subgroup are rational.

THEOREM 7. *Let* G *be a torus in characteristic* 0. *Let* Γ *be a finitely generated subgroup of* G, *and let* C *be a curve contained in* G. *If* $C \cap \Gamma$ *is infinite, then* C *is the translation of a subtorus.*

Proof: Let K be a field of finite type over $\mathbf{Q}$ such that G is isomorphic to a product of multiplicative groups over K, and such that all points of Γ are rational over K. Let $(x_1, \cdots, x_n)$ be a generic point of C over K, each x_i lying in the multiplicative group. In view of Theorem 4, C must have genus 0, and hence $K(x_1, \cdots, x_n) = K(t)$ for some parameter t. Thus each x_i is a rational function of t, say

$$x_i = \varphi_i(t) \ .$$

Let m be an integer $\geqq 1$. Infinitely many points of C must lie in some coset of Γ/Γ^m. Thus there exist elements $a_i \in K^*$ such that if we put $x_i = a_i\xi_i^m$, then the curve C' whose generic point is $(\xi_1, \cdots, \xi_n)$ over possibly a finite extension of K has infinitely many points in Γ. This implies that this curve also has genus 0. From this, we deduce that each function $\varphi_i(t)$ has at most two singularities (zero or pole). Indeed, a curve defined by an equation

$$\xi^m = a\varphi(t), \qquad a \in K^*$$

where $\varphi(t)$ is a rational function of t has genus $\geqq 1$ as soon as m is sufficiently large, and relatively prime to the order of the zeros and poles of φ, provided that φ has at least 3 singularities. We see this from the genus formula

$$2g' - 2 = -2m + \sum (e_P - 1)$$

each ramification index being equal to m in this case. Thus the term $\sum (e_P - 1)$ grows at least like $3(m - 1)$ as m tends to infinity, and the genus does become large.

Choosing our parameter t properly, we may write say for $i = 1$,

$$x_1 = a_1 t^{r_1}$$

for some integer $r_1 \neq 0$. We now contend that each x_i can be written $a_i t^{r_i}$. This will prove our theorem.

Our covering C' of C contains as an intermediate covering the curve defined by the equation

$$\xi^m = a_1 t^{r_1}(a_i\varphi_i(t))^{\pm 1} .$$

which must be of genus 0. This implies that $t^{r_1}\varphi_i(t)^{\pm 1}$ has at most two singularies, and hence clearly that $\varphi_i(t) = a_i t^{r_i}$ for some integer r_i. Our theorem is proved.

THEOREM 8. *Let K be a field of finite type over* **Q**, *and R a subring of finite type over* **Z**. *Let C be a complete non-singular curve of genus* 0, *defined over K, and let φ be a non-constant function in $K(C)$. Suppose that there are infinitely many points of C_K such that $\varphi(P)$ lies in R. Then φ has at most two distinct poles.*

Proof: Suppose φ has at least three distinct poles. We can select a model of C such that none of the poles is at infinity. This means that $K(C) = K(t)$ for some transcendental element t over K, and that

$$\varphi(t) = f(t)/g(t)$$

where f and g are polynomials in $K[t]$, $g(t)$ has leading coefficient 1 and at least three distinct roots $\alpha_1, \alpha_2, \alpha_3$. After making a finite extension of K, we may assume that these three roots are in K. Without loss of generality, we may assume that K is the quotient field of R (adjoin a finite number of elements to R) and that R is integrally closed (Proposition 2 of Chapter II, § 1). If there are infinitely many points P in C_K such that $\varphi(P) \in R$, then for such points P, the distribution $\mathfrak{a}_{\alpha_i}(P)$ as in Theorem 6 is equivalent to 0 (additive notation). This means that the function $(t - \alpha_i)/(t - \alpha_j)$, for $i \neq j$, which has no zero or pole except α_i and α_j, is such that for any value t' of t in K, $\neq \alpha_i$ or α_j, the M_R-ideal of $(t' - \alpha_i)/(t' - \alpha_j)$ lies between two fixed bounds. Excluding the primes occurring in a finite set (Proposition 5 of Chapter II, § 4) and using Theorem 5 of Chapter II, § 4 we see that $(t' - \alpha_i)/(t' - \alpha_j)$ lies in a finitely generated group.

Now we have Siegel's identity:

$$\frac{\alpha_2 - \alpha_1}{\alpha_3 - \alpha_1} \cdot \frac{t - \alpha_3}{t - \alpha_1} + \frac{\alpha_3 - \alpha_2}{\alpha_3 - \alpha_1} \cdot \frac{t - \alpha_1}{t - \alpha_2} = 1 .$$

Thus from our assumptions, we get infinitely many points in a finitely generated multiplicative group on a curve of the type considered at the beginning of this section, contradiction.

Knowing that our function φ has only two poles, we can then describe it further. On the one hand, these two poles may be both finite, determined by α, $\bar{\alpha}$ where α is a quadratic irrationality over K. Thus

$$\varphi(t) = f(t)/g(t)$$

where $g(t)$ is a power of $(t - \alpha)(t - \bar{\alpha})$, essentially a norm. One could use the decomposition theorem once more to analyze this situation. We leave it to the reader. If we make a quadratic extension of K in this case, or if all the roots of g are in K, then we can choose our parametrization of C by t in such a way that one of the poles is at infinity, and thus $\varphi(t)$ takes the form

$$\varphi(t) = f(t)/t^m$$

where $f(t)$ is a polynomial and m is an integer $\geqq 0$. The decomposition theorem again shows us that the only possibility for integral points are the obvious ones.

We could also analyze the situation first in the relative case, in order to reduce it afterwards to the absolute case or the case of number fields. Let us just state one more result.

THEOREM 9. *Let K be a function field over the constant field k of characteristic* 0, *and let R be a finitely generated subalgebra of K over k. Let C be a complete non-singular curve of genus* 0 *defined over K, and φ a non-constant function in $K(C)$. If φ has at least three distinct poles, then the points P of C_K such that $\varphi(P) \in R$ have bounded height.*

Proof: Just as in the preceding theorem, we end up with an equation $ax + by = 1$ and solutions in R^* (after possibly extending R to a larger ring). We can then proceed as before in the reduction to curves of genus $\geqq 1$ to get what we want.

We leave it to the reader to show that elements of bounded height in K^* lie in a finite number of cosets of k^*. Our theorem reduces the study of integral points on C to such points of bounded height.

Historical Note:

It was Siegel [37] who showed how one can put the Mordell-Weil theorem together with diophantine approximations (the Thue-Siegel theorem) in order to prove that a curve of genus $\geqq 1$ has only a finite number of integral points. Since at that time he did not have the strong form of Roth's theorem he had to surmount considerable difficulties in the diophantine approximations because of the lack of uniformity.

He actually worked over the ring of integers of a number field. Mahler [21] noted that by using his p-adic version of the diophantine approximations, one could allow a finite number of primes in the denominator, at any rate for curves of genus 1 over the rationals. He conjectured that the same result would hold for curves of genus $\geqq 1$ over number fields.

Once the theory of the Jacobian was sufficiently developed and once Roth's theorem was obtained, it was then natural to reconsider this question. The general version used here was presented in [18] following Siegel's (and Mahler's) method. The Jacobian replaces the theta functions, as usual, and the mechanism of the covering already used by Siegel appears here in its full formal clarity. It is striking to observe that in [15] I used the Jacobian in a formally analogous way to deal with the class field theory in function fields. In that case, Artin's reciprocity law was reduced to a formal computation in the isogeny $u \to u^{(q)} - u$ of the Jacobian. In the present case, the heart of the proof is reduced to a formal computation of heights in the isogeny $u \to mu + a$.

The decomposition theorem and the theory of distributions were created by Weil in his thesis [41], and we have followed the exposition he gave in [40], taking into account our further axiomatization of it. Siegel used it to handle curves of genus 0, but in that case made a direct analysis of the functions, which have a natural factorization, not only into divisors. In the theorems concerning curves of genus 0 we also follow Siegel [37], transporting his proofs to our more general case. The fact that Siegel's theorems for curves of genus 0 or $\geqq 1$ have a relative formulation was pointed out in [18]. An analysis of the classical proofs shows among other things that they must be separated into two parts: One, proving that certain sets of points have bounded height, and second showing what sets of bounded height look like.

As Siegel already observed, the coset argument is formally the same when dealing with curves of genus $\geqq 1$ or subgroups of multiplicative groups. The analogy between toruses and abelian varieties was again observed by Chabauty, who in two papers [5] and [6] considers the infinite intersections of a subvariety of a torus or an abelian variety with particular subgroups of finite type, namely subgroups of units and groups of rational points in a number field, respectively. Thus one is led to generalize and reformulate a conjecture of Chabauty [5] in the following manner.

Let G be a torus (resp. an abelian variety) in characteristic 0. *Let V be a subvariety of G having an infinite intersection with a subgroup of finite type Γ of G. Then V contains a finite number of translations of group subvarieties of G which contain all but a finite number of points of $V \cap \Gamma$.*

This statement was proved in § 4 for the case when V is of dimension 1 and G is a torus. One may ask whether it would not be valid also for a commutative group extension of an abelian variety by a torus.

Let us consider the statement when G is an abelian variety A and V is a curve. It then generalizes a classical conjecture of Mordell [23] to the effect that a curve of genus $\geqq 2$ over the rationals has only a finite number of rational points.

As for integral points, I would conjecture that Siegel's theorem in its absolute and relative form should be valid for abelian varieties, namely: *If A is an abelian variety in characteristic* 0, *defined over a field of finite type K, if U is an affine open subset of A, and R a finitely generated subring of K over* $\mathbf{Z}$, *then U has only a finite* *number of points in* R. In the relative case where K is function field over a constant field k of characteristic 0, and R a finitely generated subalgebra over k, the conjecture states that *integral points of* U *in* R *have bounded height, and thus lie in a finite number of cosets of* τB_k.

In this connection, the Mordell conjecture becomes a conjecture in algebraic geometry, and it is worthwhile to make further comments on it here. Let k be as above, $K = K(t)$ a function field over k, where t is the generic point of a variety T, and let $C = C_t$ be a curve of genus $\geqq 2$, defined over K. Then C_t can be viewed as a generic member of an algebraic family. The conjecture asserts that if C_t *has infinitely many rational points in* $k(t)$ (*cross sections of the parameter variety* T *in the graph of the family*), *then* C_t *is birationally equivalent over* $k(t)$ *to a curve* C_0 *defined over* k, *and all but a finite number of these points come from points of* C_0 *in* k. In particular, the graph is birationally equivalent to a product.

Evidence for this comes from the special case when $C_t = C_0$ is already defined over k, and then one obtains a classical theorem of de Franchis, to the effect that *given a variety* V *and a curve* C *of genus* $\geqq 2$ (*in characteristic* 0), *there exists only a finite number of generically surjective rational maps of* V *on* C. We give a quick proof of this theorem. Taking a generic hyperplane section of V and inducing the rational map on it, one reduces the theorem to the case when V is itself a curve. Indeed, two distinct generically surjective rational maps $f, f' : V \to C$ induce distinct generically surjective maps on the hyperplane section, as one sees by taking the induced homomorphisms on the Albanese varieties, using AV, Theorem 4 of Chapter VIII, § 2.

Assuming now that V is a curve, we have the formula for the genus

$$2g(V) - 2 = d[2g(C) - 2] + \lambda$$

where $\lambda \geqq 0$. Thus the degree of V over C is bounded. Taking suitable projective embeddings, we see that the degree of the graph of our rational maps f must be bounded. Hence these graphs Γ_f lie in finitely many algebraic families on $V \times C$. On the other hand, a generic element of such families is likewise a generically surjective rational map of V onto C (as one sees by projecting on both factors). Taking the induced homomorphisms on the Jacobians, and using the fact that an abelian variety has no algebraic family of abelian subvarieties, we see that all induced maps coming from the same family differ by translations. We now use the fact that C is not equal to a non-zero translation of itself in its Jacobian. We conclude that a graph Γ_f actually must constitute by itself a maximal algebraic family on $V \times C$, and finally that there is only a finite number of such graphs, or maps f. This concludes the proof. (When V is a curve, we do not need characteristic 0, only the assumption that the map $f: V \to C$ is separable, to be able to use the genus formula above.)

The Mordell conjecture thus gives rise to diophantine criteria for lowering fields of definition, and we actually proved such a criterion in the context of integral points (Proposition 2).

Finally, let us mention Chabauty's partial result. He proves the following statement in [6]: *Let C be a complete non-singular curve of genus $\geqq 2$ defined over a $\mathfrak{p}$-adic field K. Let C be embedded in its Jacobian over K, and let Γ be a finitely generated subgroup of J_K. If the rank of Γ is smaller than* (*strictly*) *the genus of C, then $C \cap \Gamma$ is finite.* We sketch his argument. An infinite number would have a point of accumulation, say P_0. In a neighborhood of P_0, there will be infinitely many points of $\Gamma \cap C$, and thus without loss of generality (by the non-archimedean behaviour) we may assume that Γ itself lies in such a neighoborhood. The assumption on the rank would then imply a linear relation on the differentials of the first kind on C or on J (those of C being induced by those of J), which is impossible.

CHAPTER VIII

Hilbert's Irreducibility Theorem

Let k be a field, and $f(t_1, \cdots, t_r, X_1, \cdots, X_s)$ a polynomial with coefficients in k which is irreducible as a polynomial in $r+s$ variables. We ask whether there exist values (t') of (t) in k such that the polynomial $f(t', X)$ is irreducible as a polynomial in (X) with coefficients in k. We shall exhibit a large variety of fields for which this is true.

Suppose $f(t, X)$ has coefficients in $k(t)$, i.e. is in $k(t)[X]$, and assume it is irreducible as a polynomial in (X) over $k(t)$. Multiplying f by a polynomial $\phi(t)$, to make the coefficients lie in $k[t]$ does not change this irreducibility property. If we then divide the resulting polynomial by the greatest common divisor of its coefficients, we obtain a polynomial $f_1(t, X) \in k[t, X]$ which is irreducible over k. Conversely, if $f(t, X)$ is in $k[t, X]$ and irreducible over k, then it is irreducible in $k(t)[X]$.

Given $f(t, X) \in k(t)[X]$, we denote by $U_{f,k}$ (or simply U_f if the reference to k is clear) the subset of the affine space S_k^r consisting of those points $(t_1', \cdots, t_r')$, with $t_i' \in k$, at which the coefficients of f are defined and such that $f(t', X)$ is irreducible in $k[X]$ over k. Such $U_{f,k}$ will be called *basic Hilbert sets.* The intersection of a finite number of basic Hilbert sets with a finite number of non-empty Zariski open subsets of S_k^r will be called a *Hilbert subset* of S_k^r. If $r=1$, then $S_k^1 = k$ and the non-emply Zariski open sets are the complements of finite subsets of k.

A field k is called *Hilbertian* if the Hilbert subsets of S_k^r are not empty (and thus are infinite). If R is a subset of k, we shall say that R is *Hilbertian in* k if every Hilbert subset of S_k^r contains some element $(t_1', \cdots, t_r')$ with $t_i' \in R$. A ring will be said to be *Hilbertian* if it is Hilbertian in its quotient field.

Among other results, we shall prove in this chapter that all rings of finite type over the prime ring, and of transcendence degree $\geqq 1$ in characteristic >0 are Hilbertian. In particular, every number field is Hilbertian. The proof of this fact will depend only on elementary calculus.

It will be shown also that the essential case occurs when $r=1$

and $s = 1$, i.e. when there is one parameter and one variable. The general case can be reduced to this one by simple formal arguments. The essential case will be seen to be equivalent to the lifting of rational points in coverings of curves. From this point of view one is thus led to diophantine questions.

In § 6 we shall apply the irreducibility property to abelian varieties, and show for instance that if an abelian variety or a curve of genus $\geqq 1$ is defined over a field of finite type over $\mathbf{Q}$, then there exist infinitely many non-degenerate specializations of it into a number field which induce an injection on the sets of rational points.

We make one final observation, as Hilbert does [12], concerning applications to Galois theory. Suppose that we can construct a purely transcendental extension $k(t_1, \cdots, t_r)$ of k and a finite Galois extension E of $k(t_1, \cdots, t_r)$ with group G. If k is Hilbertian then we can specialize the (t) on (t') in k in such a way that the decomposition group remains of the same order as that of G, and thus obtain a Galois extension of k with group G. This can be done in particular with the symmetric groups over any field k, thereby allowing either the possibility of constructing a Galois extension with the symmetric group if k is Hilbertian or of proving that a field is not Hilbertian (for instance, a $\mathfrak{p}$-adic field). However, it does not seem to be known whether the power series ring in several variables over a field (or a unique factorization domain with infinitely many primes) is Hilbertian.

§ 1. *Irreducibility and integral points*

We consider the case of one t and one X. Let us write

$$f(t, X) = a_n(t)X^n + \cdots + a_0(t)$$

with $a_i(t) \in k[t]$. In studying special values of t in k, we always *exclude* the finite number which make a_n vanish.

Suppose that $f(t, X)$ is irreducible over $k(t)$. If f has a factorization

$$f(X) = a_n(t) \prod_{i=1}^{n} (X - \alpha_i)$$

in the algebraic closure of $k(t)$, every value t_0 of t in k determines a homomorphism

$$k[t] \to k[t_0] = k$$

which can be extended to the ring generated by the roots $\alpha_1, \cdots, \alpha_n$ of f (under our assumption that $a_n(t_0) \neq 0$). The image of these roots is determined up to a permutation. Let α_i' $(i = 1, \cdots, n)$ be these images. If we have a factorization

$$f(t_0, X) = g_0(X)h_0(X)$$

in $k[X]$, the coefficients of g_0 and h_0 are polynomial functions of the α_i'. Let g and h be the polynomials corresponding to these functions, giving a factorization

$$f(t, X) = g(X)h(X)$$

in the algebraic closure of $k(t)$. Since f is assumed to be irreducible, one of the coefficients of g or h does not lie in $k(t)$, say y. Then the ring $k[t, y]$ is the affine ring of a curve C (which may be reducible) over k. The factorization $f(t_0, X) = g_0h_0$ gives rise to a rational point (t_0, y_0) on our curve C which has projection t_0 on the first factor.

If one writes $f(t, X) = g(X)h(X)$ in all possible manners in the algebraic closure of $k(t)$, with g and h of degree $\geqq 1$, there will be each time a coefficient which does not lie in $k(t)$, whence we get a finite number of curves $C_1, \cdots, C_m$ as above. Each value t_0 in k which cannot be lifted to a rational point of any of these curves will be such that $f(t_0, X)$ is irreducible in $k[X]$.

If k is the quotient field of a ring R, and if we replace y above by $\varphi(t)y$ for a suitable polynomial $\varphi(t) \in R[t]$, we may suppose that y is integral over $R[t]$. Excluding the finite number of points t_0 for which $\varphi(t_0) = 0$, one sees that y_0 and $\varphi(t_0)y_0$ are rational or irrational over k simultaneously. If, however, they are in k, then $\varphi(t_0)y_0$ will be integral over R, and hence in R if R is integrally closed (as is usual in practice).

By an *affine plane curve* C over a field k, we shall mean an affine curve whose ring is of type $k[t, y]$ over k, with t transcendental over k. We do not assume that C is absolutely irreducible, so that $k(t, y)$ need not be regular over k. If R is a subset of k, we denote by $U_{t,R}(C)$ or simply $U_{t,R}$ the set of elements $t_0 \in R$ such that there is no point $P \in C_k$ for which $t(P) = t_0$, viewing t as a function on C. From our preceding discussion, we have:

Proposition 1. *Let k be a field, and R a subset of k. Every Hilbert subset of k contains a finite intersection of sets of type $U_{t,R}(C)$ for a finite number of affine plane curves C over k.*

The Hibert subsets of k mentioned here are those determined by irreducible polynomials $f(t, X)$ with one t and one X. We shall see in § 3 that the restriction to one X can be dispensed with.

Let us consider an affine plane curve over k, as above, with y not in $k(t)$. If y is not separable over $k(t)$, then a power y^{p^m} will be for m large and $p =$ characteristic. If y^{p^m} actually lies in $k(t)$, then to find elements of $U_{t,R}$ one analyzes k^*/k^{*p} directly. Let us exclude this case. We are therefore reduced to studying the case where y is separable over $k(t)$ but not in $k(t)$.

Let k_1 be the algebraic closure of k in $k(t, y)$. Then (t, y) determines an absolutely irreducible curve C_1 defined over k_1 having a conjugate C_1^σ over k not equal to C_1 if $k \neq k_1$. Every rational point of C in k must lie on both C_1 and C_1^σ, and hence there can be only a finite number of such points if C is not absolutely irreducible under the hypothesis that y is separable over $k(t)$. In the sequel, in trying to find points in $U_{t,R}$, we can therefore assume that C is absolutely irreducible.

Suppose that we have a variety C of dimension 1 over k and a non-constant function t in $k(C)$. If for instance k is of finite type over $\mathbf{Q}$, R of finite type over $\mathbf{Z}$, and C of genus $\geqq 1$ then we know that there is only a finite number of points P in C_k such that $t(P) \in R$. The worst case would therefore occur when C has genus 0, i.e. is a pure curve.

In that case, following Siegel, one can apply the analysis of curves of genus 0 to estimate the number of bad points, but using an idea of Néron, it is just as well to reduce the general case to that of curves of genus $\geqq 1$, by means of the following trick.

We start with our plane curve C, with generic point (t, y) over k. If we write $t = f(u)$ for a suitable polynomial $f(u)$ in $k[u]$, then we get a new curve C' with generic point (t, y, u) over k, which will have genus $\geqq 1$ (and in fact arbitrarily high for suitable f). We now look for special values u_0 which cannot be lifted to rational points (t_0, y_0, u_0) of C' in k. Of course, $t_0 = f(u_0)$ is determined, and if f is taken to have coefficients in a subring R of k, and u_0 is selected in R, then t_0 will also lie in R, so that our integrality conditions are preserved.

In order to find polynomials f we simply use the genus formula for a covering. We have two cases:

Case 1. *The function t has at least* 3 *singularities* (*zeros or poles*) *on the complete non-singular curve determined by C.*

We then take m large, relatively prime to the characteristic of k, to the degree $[k(C): k(t)]$, and to the orders of zeros and poles of t. Then the curve C' whose function field is $k(C)(u)$ over k is absolutely irreducible, and of large genus, because the formula

$$2g' - 2 = -2m + \sum(e_j - 1)$$

shows that the term $\sum(e_j - 1)$ will grow at least like $3(m - 1)$, each e_j being equal to m.

Case 2. *The function t has exactly* 2 *singularities.*

Then we can find a function x in $k(C)$ such that $k(C) = k(x)$, and $t = x^r$ with an integer $r > 0$, prime to the characteristic p of k, or $t = x^{p^n} - a$ for some $a \in k$, and an integer $n \geqq 0$. If $k(C) \neq k(t)$ then $r > 1$ or $n > 0$. Leaving aside the characteristic difficulties, and looking only at $t = x^r$, we take for $f(u)$ any polynomial of degree m prime to p with m distinct roots to see that the genus of C' again becomes large if m becomes large.

We have thus shown that any subring R of a number field is Hilbertian. In the next section, we shall given an elementary proof of this fact.

In characteristic $p > 0$, we do not have the finiteness of integral points because of the Frobenius automorphism. However it suffices to use the Mordell-Weil theorem, which gives us a sufficiently strong way of counting points, to see that they are Hilbertian. In fact, having obtained our curve of genus $\geqq 1$, we embed it in its Jacobian over k. If J_k is finitely generated, then we found an upper bound for the number of points of height $\leqq B$ in Chapter V, § 4. In practice, if R is a subring of k, the number of points of R of height $\leqq B$ is very much larger than a power of $\log B$, being in fact of the order of magnitude of B itself (for instance for the ordinary integers $\mathbf{Z}$ and, as the reader himself will verify immediately, for a polynomial ring in one variable over a finite constant field). Thus if we have a theory of heights on k, if $f: C' \to \mathbf{P}^1$ is a generically surjective rational map of C' on $\mathbf{P}^1$ defined over k, and if C' has genus $\geqq 1$, then we know that for a point $P' \in C'_k$ we have

$$h(f(P')) \approx h(P')^d$$

where d is the degree of f (Property 4 of heights, which we are applying in a case where it could be verified more trivially).

Let us look at polynomials over a finite field. Let $R=k_0[w]$ be a polynomial ring in one variable over the finite field k_0 with q elements. The number of elements of R of degree $<r$ is q^r and one also sees that we could replace thc degree by the logarithmic height, computed with respect to the obvious model. Let C be an affine plane curve with generic point (t, y) over the quotient field k of R. From the above discussion, we see that if y is separable over $k(t)$, then there exists a number α, with $0<\alpha<1$ such that the number of elements of $U_{t,R}(C)$ of degree $<r$ is

$$q^r + O(q^{\alpha r}) \qquad \text{for } r \to \infty .$$

If y is not separable, then one must examine this case more closely. We shall deal with it in connection with arbitrary ground fields k in § 4.

§ 2. *Irreducibility over the rational numbers*

We consider the concrete case of the rational numbers, or rather the ring of integers $\mathbf{Z}$. Let (t, y) be a generic point of an affine plane curve C over $\mathbf{Q}$, and suppose $y \notin \mathbf{Q}(t)$. After multiplying y by a suitable polynomial in $\mathbf{Z}[t]$, we may assume that y is integral over $\mathbf{Z}[t]$. Thus if any value t_0 (excluding a finite number) in $\mathbf{Q}$ extends to a point (t_0, y_0) on C with $y_0 \in \mathbf{Q}$, then y_0 lies in $\mathbf{Z}$ if t_0 lies in $\mathbf{Z}$.

We may view y as an algebraic function of t over the reals, and as such it has an expansion at infinity:

$$y = \varphi(t) = at^{n/e} + \cdots + b + c\frac{1}{t^{1/e}} + \cdots$$

with coefficients $a, b, c, \cdots$ that are complex numbers, and $t^{1/e}$ one of the e-th roots of t taken as parameter. We choose in fact $t^{1/e}$ to be real. In that case, if there exist infinitely many values of t tending to infinity in $\mathbf{R}$ such that $\varphi(t)$ is real, then the coefficients are in fact real. Indeed, if one of them were complex, let us consider the one most to the left, say ξ. The angle of the term having ξ for its coefficient is $\neq 0, \pi$. As t tends to infinity, the term with this coefficient dominates the series to the right of it. Hence there cannot be any cancellations, and for $t \to \infty$ the values $\varphi(t)$ would be complex, contradicting the hypothesis.

We now study the integral points on the analytic curve $(t, \varphi(t))$ and prove the following theorem, concerning their distribution.

THEOREM 1. *Let $\varphi(t)$ be a function of a real variable having a power series expansion*

$$\varphi(t) = at^{n/e} + \cdots + b + c\frac{1}{t^{1/e}} + \cdots$$

with coefficients $a, b, c, \cdots$ in $\mathbf{R}$, in terms of the real parameter $t^{1/e}$, and converging for all sufficiently large values of t. Assume that $\varphi(t)$ does not lie in $\mathbf{R}[t]$, i.e. is not a polynomial in t. Suppose that that there is an infinity of positive integers

$$t_0 < t_1 < t_2 < \cdots$$

such that $\varphi(t_i)$ is in $\mathbf{Z}$. Then there exists an integer i_0, an integer $m > 0$, and a real number $\lambda > 0$, such that for all $i > i_0$ we have

$$t_{i+m} - t_i > t_i^{\lambda}\,.$$

In other words, the t_i become more widely spaced. As a corollary, we get an estimate which we can handle better in practice.

COROLLARY 1. *Let $\varphi(t)$ be as in the theorem. There exists a real number α with $0 < \alpha < 1$ such that the number $N(B)$ of $t_i \leqq B$ for which $\varphi(t_i)$ is in $\mathbf{Z}$ satisfies the inequality*

$$N(B) \leqq B^{\alpha}$$

for all B sufficiently large.

Proof. Choose $0 < \beta < 1$. Let N_1 be the number of integers t_i such that $t_i \leqq B^{\beta} + 1$, and N_2 the number of t_i such that $B^{\beta} < t_i \leqq B$. Write $N_2 = sm + m_0$ with $s \geqq 0$ and $0 \leqq m_0 < m$. By the Theorem one verifies immediately, using a crude estimate, that

$$B^{\beta} + sB^{\beta\lambda} \leqq B\,,$$

whence $s \leqq B^{1-\beta\lambda}$ and thus

$$N(B) \leqq N_1 + N_2 \leqq B^{\beta} + 1 + m_0 + B^{1-\beta\lambda}\,.$$

From this one immediately deduces the existence of our number α.

COROLLARY 2. *Let U be a Hilbert subset of $\mathbf{Q}$. Then there exists a number α, with $0 < \alpha < 1$, such that the number of positive integers $\leqq B$ lying in U is at least equal to*

$$B - B^{\alpha}$$

for all B sufficiently large.

Proof. Apply Proposition 1 of § 1.

COROLLARY 3. *A Hilbert subset of* **Q** *contains infinitely many prime numbers.*

Proof. Apply the prime number theorem.

COROLLARY 4. *A Hilbert subset of* **Q** *is dense for the ordinary topology and every p-adic topology on* **Q**.

Proof. If $f(t, X)$ is irreducible over $\mathbf{Q}(t)$, then so is $f(a + 1/t, X)$. Integral values of t tending to infinity and making the specialized polynomial irreducible will be such that $a + 1/t$ is close to a for the ordinary topology. For the p-adic topology, one takes $f(a + tp^{\nu}, X)$ with ν large.

In order to prove Theorem 1, we need a mean value theorem due to H. A. Schwarz.

LEMMA. *Let $\varphi(t)$ be a function m times continuously differentiable in the interval $t_i \leqq t \leqq t_{i+m}$. We suppose that $t_i < t_{i+1} < \cdots < t_{i+m}$ are real numbers (not necessarily in* **Z**). *Then there exists a number τ with $t_i < \tau < t_{i+m}$ such that*

$$\frac{\varphi^{(m)}(\tau)}{m!} = \frac{\begin{vmatrix} 1 & t_i & t_i^2 & \cdots & t_i^{m-1} & \varphi(t_i) \\ & \cdots & & \cdots & & \cdots \\ 1 & t_{i+m} & t_{i+m}^2 & \cdots & t_{i+m}^{m-1} & \varphi(t_{i+m}) \end{vmatrix}}{V_m}$$

where V_m is the Vandermonde determinant.

(We see that the numerator differs from V_m only by the last column).

Proof. I owe this proof to Walter Strodt. Let $F(t)$ be the function

$$F(t) = \begin{vmatrix} 1 & t_i & t_i^2 & t_i^{m-1} & \varphi(t_i) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & t_{i+m-1} & t_{i+m-1}^2 & \cdots & \varphi(t_{i+m-1}) \\ 1 & t & t^2 & t^{m-1} & \varphi(t) \end{vmatrix}$$

Then $F(t)$ is 0 when t is equal to the m points $t_i, \cdots, t_{i+m-1}$. For some constant C the function

$$G(t) = F(t) - C(t - t_i)(t - t_{i+1}) \cdots (t - t_{i+m-1})$$

will also vanish at t_{i+m}, that is to say at $m + 1$ points. Hence $G^{(m)}(t)$ vanishes at least at one point τ between t_i and t_{i+m}. But $G^{(m)}(t) = F^{(m)}(t) - m!\,C$ (all the other terms will cancel). From this we conclude that

$$F^{(m)}(\tau) = m!\,C\,.$$

One sees immediately that $F^{(m)}(t)$ will have a 0 everywhere in the bottom line of its matrix, except for the last term, which will be $\varphi^{(m)}(\tau)$. Hence

$$F^{(m)}(\tau) = \varphi^{(m)}(\tau) V_{m-1}$$

where V_{m-1} is the small Vandermonde determinant. Since

$$C = \frac{F(t_{i+m})}{(t_{i+m} - t_i) \cdots (t_{i+m} - t_{i+m-1})}$$

and since $F(t_{i+m})$ is none other than the numerator of the expression in the statement of the lemma, we see that our lemma is proved, taking into account the fact that the product of V_{m-1} and the denominator of C is precisely equal to V_m.

Let us apply our lemma, taking for m an integer such that $\varphi^{(m)}(t)$ has no negative powers of $1/t^{1/e}$, hence its development is of the type

$$\varphi^{(m)}(t) = \xi \frac{1}{t^{\lambda}} + \cdots$$

with ξ real. Since φ is not a polynomial in t, its derivative $\varphi^{(m)}$ is not identically 0, and so we may assume $\xi \neq 0$ and $\lambda > 0$. This term dominates the series for t large. Hence for sufficiently large i, we have $\varphi^{(m)}(\tau) \neq 0$. Observe that the determinant in the numerator of the expression in the statement of the lemma is an integer. We have just seen that $\varphi^{(m)}(\tau)$, and a fortiori $\varphi^{(m)}(\tau)/m!$ is small, of order of magnitude $t^{-\lambda}$. This implies that V_m is large. But V_m is a product of differences of the t_j, and we can replace each such difference by $t_{i+m} - t_i$. This makes V_m only larger. There will be $m(m + 1)/2$ such terms. We find therefore

$$t_{i+m} - t_i > t_i^{\lambda'}$$

with a suitable λ', taking an $m(m + 1)/2$-root.

§ 3. *Reduction steps*

We study systematically the following reductions of the Hilbert polynomials to the case of one t, several X, several t and several X, and finite separable extensions.

a. One t, several X.

Let k be a field, and $f(X_1, \cdots, X_s)$ a polynominal with coefficients in k, of degree $<d$ in each X_i. We denote the set of such polynomials by $P(s, d, k)$. *Kronecker's specialization* S_d transforms f into a polynomial in 1 variable, namely

$$S_d f(Y) = f(Y, Y^d, \cdots, Y^{d^{s-1}}) .$$

If we have a monomial $X_1^{i_1} \cdots X_s^{i_s}$, then S_d applied to this monomial yields

$$Y^{i_1 + i_2 d + \cdots + i_{s-1} d^{s-1}} .$$

Taking into account the d-adic expansion of an integer $\geqq 0$, we see S_d is a bijection of $P(s, d, k)$ on the polynomials in $k[Y]$ of degree $\leqq d^s - 1$. Furthermore, if f, g and fg are in $P(s, d, k)$ then

$$S_d(fg) = S_d(f) S_d(g) .$$

Let $f \in P(s, d, k)$ and suppose that $S_d f$ is reducible:

$$S_d f(Y) = G(Y) H(Y) .$$

Then the degree of G and H is $\leqq d^s - 1$ and hence $G = S_d g$ and $H = S_d h$ for some g, h uniquely determined by G, H respectively. From this we deduce *Kronecker's criterion*:

f is irreducible in k if and only if for every factorization $S_d f = GH$ with $G = S_d g$ and $H = S_d h$, the polynomial gh contains a monomial

$$\varphi X_1^{i_1} \cdots X_s^{i_s}$$

with $\varphi \in k$, $\varphi \neq 0$, of degree $\geqq d$ for one of the X_i.

PROPOSITION 2. *Let k be a field, $f(t, X_1, \cdots, X_s)$ a polynomial with coefficients in k, of degree $<d$ in each X_i, and let S_d be Kronecker's substitution. Let*

$$S_d f = \prod g_j(t, Y)$$

be the factorization of $S_d f$ in $k(t)[Y]$ into irreducible factors in Y. Then, except for a finite number of values of t_0 in k, every t_0 in k

such that each $g_j(t_0, Y)$ *is irreducible in* $k[Y]$ *is also such that* $f(t_0, X_1, \cdots, X_n) = f_0$ *is irreducible in* $k[X]$.

Proof. Except for a finite number of t_0, we have

$$S_d f_0 = \prod g_j(t_0, Y)$$

a factorization of $S_d f_0$ into irreducible factors in $k[Y]$. Excluding still another finite set of t_0 (those for which some coefficient $\varphi(t)$ in a monomial in Kronecker's criterion will vanish), it is clear that f_0 will be irreducible.

In particular, we see that every Hilbert subset of k determined by a polynomial $f(t, X_1, \cdots, X_s)$ contains a Hilbert set determined by a finite number of polynomials $g_j(t, Y)$ with a single Y.

b. Several t, several X

We consider a polynomial $f(t_1, \cdots, t_r, X_1, \cdots, X_s)$ as a polynomial in t_1 and $r + s - 1$ variables $t_2, \cdots, t_r, X_1, \cdots, X_s$. We specialize t_1 to preserve irreducibility, and continue inductively. Thus on the Hilbert set $U_{f,k}$ in S_k^r we get a fibering, and we can determine its points by succesive consideration of the preceding case.

c. Finite separable extensions

We shall prove that if k is Hilbertian, then any finite separable extension of k is also Hilbertian. We prove in fact a stronger result:

PROPOSITION 3. *Let E be a finite separable extension of a field k. Let $f(t, X) \in E(t)[X]$ be irreducible over $E(t)$. Assume that the leadind coefficient of f in X is* 1, *and that if σ ranges over the distinct isomorphisms of E over k, then the conjugates f^σ of f are distinct Put*

$$F(t, X) = \prod_\sigma f^\sigma(t, X) .$$

Then $F(t, X)$ is in $k(t)[X]$ and is irreducible. Excluding a finite number of $t_0 \in k$, $F(t_0, X)$ is irreducible over k if and only if $f(t_0, X)$ is irreducible over E.

Proof. By Galois theory, $F(t, X)$ is in $k(t)[X]$. If we had $F(t, X) = G(X)H(X)$, a factorization over $k(t)$, then we could assume G and H have leading coefficient 1. We conclude that $f(t, X)$ divides G or H over $E(t)[X]$, say G. But then f^σ also divides G

for each σ. As the f^σ are supposed to be distinct, their product divides G, a contradiction unless $F = G$. We now choose t_0 in k such that $F(t_0, X)$ is irreducible over k. Then it is clear that $f(t_0, X)$ will be irreducible over E, taking into account unique factorization in $E[X]$.

To apply the proposition even to the case where the conjugates f^σ might not be distinct, we make a simple transformation on f. Indeed, there exists an element $\varphi = \varphi(t) \in E(t)$ such that if we put

$$g(t, X) = f(t, X + \varphi)$$

all the g^σ are distinct. For instance, the constant term of g is $f(t, \varphi)$ and the existence of the desired φ is obvious. We can apply Proposition 3 to g, whose irreducibility over $E(t)$ is equivalent to that of $f(t, X)$. In particular, we get:

Corollary. *Let E be a finite separable extension of a field k. Every Hilbert subset of E contains a Hilbert subset of k.*

For a treatment of the purely inseparable case the reader may consult [13], where the analogous result is proved.

§ 4. *Function fields*

In dealing with function fields, one is led by induction to the cases of purely transcendental extensions and finite extensions. Concerning the first, we show that Hilbert sets contain an abundance of elements.

Proposition. *Let k be an infinite field and w transcendental over k. Let t be transcendental over $k(w)$, and y separable over $k(w, t)$ but not in $k(w, t)$. Let U be the Hilbert subset of $k(w)$ determined by (t, y) as in § 1. Let m be an integer $\geqq 1$. In the (w, t) plane, there exists a non-empty Zariski open set V such that for any $\lambda_0 \in k$ (except possibly a finite number depending on U, m, V) all the values t_0 of t of type*

$$t_0 = \tau_0 + \lambda_0(w - w_0)^m$$

with τ_0, $w_0 \in k$ and $(\tau_0, w_0) \in V$ are in U.

Proof. After multiplying y by a polynomial in $k[w, t]$ we may assume that y is integral over $k[w, t]$. We consider the surface with generic point (w, t, y) over k, which can be viewed as a covering of the plane (w, t).

As we assume y separable over $k(w, t)$, there exists a relation

$$\varphi(w, t) = P(w, t, Y)f(w, t, Y) + Q(w, t, Y)f'(w, t, Y)$$

where $f(w, t, Y)$ is the irreducible polynomial of y over $k[w, t]$, f' its derivative with respect to Y, and P, Q are in $k[w, t, Y]$ while $\varphi \in k[w, t]$, $\varphi \neq 0$. Consequently, if (w_0, τ_0, y_0) is a point of the surface in k such that $\varphi(w_0, \tau_0) \neq 0$, then y has an expansion in formal power series:

$$y = \sum c_{ij}(t - \tau_0)^i (w - w_0)^j$$

with $c_{ij} \in k$. (Geometrically speaking, the covering is unramified.)

The zeros of $\varphi(w, t)$ form a closed Zariski set, whose complement is the open set V. From now on, we assume that (w_0, τ_0) is in V and $w_0, \tau_0 \in k$.

If there does not exist a point (w_0, τ_0, y_0) on the surface in k above (w_0, τ_0), then for all $\lambda_0 \in k$, there is no point on the (t, y)-curve in $k(w)$ above

$$\tau_0 + \lambda_0(w - w_0)^m.$$

Indeed, such a point would have polynomials in $k[w]$ as coordinates, and the homomorphism of this ring mapping w on w_0 would then give a point of the surface in k above (w_0, τ_0).

Suppose, on the other hand, that (w_0, t_0, y_0) is a point of the surface in k. If we substitute

$$\lambda(w - w_0)^m$$

for $(t - \tau_0)$ in the power series, we obtain a homomorphism of $k[w, t, y]$ into the power series ring $k[[w - w_0]]$. We have to choose λ in such a way that y does not map into $k[w - w_0] = k[w]$ (it will in any case be integral over $k[w]$). In other words, the series must not stop.

Since y does not lie in $k[w, t]$, infinitely many c_{ij} are not equal to 0. We can write

$$\begin{aligned} y_{\lambda, m} &= \sum_{i,j} c_{ij}\lambda^i (w - w_0)^{mi+j} \\ &= \sum_{\nu} c_\nu(\lambda)(w - w_0)^\nu . \end{aligned}$$

Each $c_\nu(\lambda)$ is a polynomial in λ.

If, for some value λ_0 of λ in k, all but a finite number of $c_\nu(\lambda_0)$

are 0, then $y_{\lambda_0,m}$ considered as a rational function of w has a pole at infinity ($w = \infty$) of order equal to the largest ν such that $c_\nu(\lambda_0) \neq 0$. But we have

$$f(w, \tau_0 + \lambda_0(w - w_0)^m, y_{\lambda_0,m}) = 0 .$$

Since m is fixed, still excluding a finite number of λ_0, we see that the order of a pole of $y_{\lambda_0,m}$ at infinity is bounded (as a function of the coefficients of f, which is fixed throughout). If we choose ν_1 larger than this critical value and such that $c_{\nu_1}(\lambda_0) \neq 0$ (which is possible since infinitely many $c_{ij} \neq 0$) we get a $y_{\lambda_0,m}$ whose series cannot stop, because if it did, we would get a bad pole of $y_{\lambda_0,m}$. This concludes the proof.

In case y is not separable over $k(w, t)$, then either y^{p^r} is not in $k(w, t)$ for any r, in which case it becomes separable for r large enough and we can apply our proposition to it; or y^{p^r} does lie in $k(w, t)$ for some r, so that we may write

$$y^{p^r} = \varphi(w, t)$$

with $\varphi(w, t) \in k[w, t]$ and we may assume that $\varphi(w, t)$ is not a p-th power in $k[w, t]$. We distinguish two cases: the first, where there occurs some power of t or w in φ which is not divisible by p, and the second, where p divides the exponent of t and w in each term of φ. In the second case, some coefficient of φ in k is not a p-th power.

In the first case, take m greater than the degree of φ. Then setting $t = \lambda_0 w^m$ for any $\lambda_0 \in k$, $\lambda_0 \neq 0$, the polynomial

$$\varphi(w, \lambda_0 w^m)$$

will have the same set of non-zero coefficients as φ (this is the Kronecker substitution). If

$$\varphi(w, t) = \sum c_{ij} w^i t^j$$

then

$$\varphi(w, \lambda_0 w^m) = \sum c_{ij} \lambda_0^j w^{i+mj} .$$

If one of the $e_{ij} \neq 0$, and p does not divide i or j, then we simply pick m so that p does not divide $i + mj$.

The same argument could be applied after making an arbitrary translation (w_0, τ_0) on (w, t).

In the second case, when p divides each i and j, then we can choose any m greater than the degree of φ and again any $\lambda_0 \in k$. If c_{ij} is not a p-th power in k, then $c_{ij}\lambda_0^j$ will also not be a p-th power and hence $\varphi(w, \lambda_0 w^m)$ will not be a p-th power in $k[w]$.

This takes care of the remaining case in our treatment of polynomial rings over finite fields. Thus, up to a purely inseparable extension (left open in our third reduction step) we have proved the following theorem.

Theorem 2. *Let R_0 be a ring, and R a finitely generated ring over R_0. Let k_0 and k be their quotient fields respectively. Then if $R_0 = \mathbf{Z}$, or if the transcendence degree of k over k_0 is $\geqq 1$, R is Hilbertian.*

Actually if R_0 is a finite field, k is separable over k_0, and hence our proof in this case is complete.

§ 5. *Abstract definition of Hilbert sets*

It will be convenient to reformulate the definition of Hilbert sets in the context of general rings. As usual, rings are without zero-divisors.

Let R be a ring. As we know, spec (R) has the Zariski topology, a basis for the open sets being the subsets spec(R_ξ) with $\xi \in R$, $\xi \neq 0$, and spec(R_ξ) is identified with the subset of prime ideals $\mathfrak{p}$ of spec(R) which do not contain ξ. These spec(R_ξ) will be called *basic Zariski sets.*

Let S be a finitely generated ring over R. We shall say that it is *algebraic* over R if all its elements are algebraic over the quotient field $K(R)$ of R. Its quotient field $K(S)$ is then a finite extension of $K(R)$. Assume that S is such a ring. The *basic Hilbert set associated with* S is the subset U_S, or $U_{S/R}$, of spec(R), consisting of those $\mathfrak{p}$ having the following two properties:

(i) There is exactly one point $\mathfrak{P}$ in spec(S) lying above $\mathfrak{p}$.

(ii) If we denote by $k(\mathfrak{p})$ and $k(\mathfrak{P})$ the residue classs fields of these points, then

$$[k(\mathfrak{P}) : k(\mathfrak{p})] = [K(S) : K(R)] .$$

By a *Hilbert subset of spec*(R) we shall mean the intersection of a finite number of basic Zariski sets with a finite number of basic Hilbert sets.

If S is of the above type, then for some $\xi \in R$, $\xi \neq 0$, S is integral over $R_\xi = R[1/\xi]$, and thus above each $\mathfrak{p} \in \operatorname{spec}(R_\xi)$ there exists at least one $\mathfrak{P}$ in $\operatorname{spec}(S)$. As we are interested only in intersections of subsets of $\operatorname{spec}(R)$ with basic Zariski sets, we may in practice assume that S is *finite* over R, i.e. integral and finitely generated over R.

Let us recall some facts from commutative algebra. If R is integrally closed, and S integral over R, and if $\mathfrak{p} \in \operatorname{spec}(R)$ and $\mathfrak{P} \in \operatorname{spec}(S)$ lies above $\mathfrak{p}$, then the separable degree

$$[k(\mathfrak{P}) : k(\mathfrak{p})]_s$$

is at most equal to $[K(S) : K(R)]$. There are cases where this does not hold true for the degree. However, if R is a *regular* ring (i.e. one for which each local ring $R_\mathfrak{p}$ is regular) then we do have

$$[k(\mathfrak{P}) : k(\mathfrak{p})] \leqq [K(S) : K(R)] \,.$$

(This could be proved for instance by "blowing up" the given point $\mathfrak{p}$, i.e. taking a suitable discrete valuation ring $\mathfrak{o}$ of K, whose maximal ideal $\mathfrak{m}$ restricts to $\mathfrak{p}$, and whose residue class field $\mathfrak{o}/\mathfrak{m}$ is purely transcendental over $k(\mathfrak{p})$. The point $\mathfrak{P}$ in $\operatorname{spec}(S)$ would lift to a point in the integral closure of $\mathfrak{o}$ in $K(S)$, and we could apply the known theorem for valuation rings.)

In practice, when we deal with rings R, they will be parameter rings, that is, we deal with schemes over R, and there will always exist an element $\xi \in R$, $\xi \neq 0$, such that R_ξ is regular (the simple points on a variety form a Zariski open set).

In order to ensure the validity of the conclusion of the next proposition, we see from our preceding discussion that we must make some restrictions on our rings. However, with such assumptions, guaranteeing the degree inequality, the proof is obvious.

Proposition 5. *Let $S' \supset S \supset R$ be three rings, with S, S' finite over R, and R, S integrally closed. Let $\mathfrak{p} \in U_{S'/R}$, and let $\mathfrak{P}$, $\mathfrak{P}'$ be the points lying above it in spec(S) and spec(S') respectively. Assume that $S_\mathfrak{p}$ and $R_\mathfrak{p}$ are regular, or that $k(\mathfrak{P}')$ is a separable extension of $k(\mathfrak{p})$. Then $\mathfrak{P}$ is in $U_{S'/S}$, and $\mathfrak{p}$ is in $U_{S/R}$.*

In our applications, $R = k[t_1, \cdots, t_r]$ is a finitely generated ring over a field k, with $t_1, \cdots, t_r$ algebraically independent over k. If, as in the proposition, we are concerned with a finite extension S' of a ring S, then using the Noether normalization theorem, one can find a polynomial ring R so that the search of points in $U_{S'/S}$

is reduced to that of points in $U_{S'/R}$.

Over the ring $k[t_1, \cdots, t_r]$ we can relate the sets U_f of § 1 to our sets $U_{S/R}$ of the previous discussion:

PROPOSITION 6. *Let k be a field, $k[t_1, \cdots, t_r]$ a purely transcendental ring extension R of k. Let S be finite over R. Then there exists an irreducible polynomial $f(t_1, \cdots, t_r, X)$ such that $U_{S/R}$ contains $U_{f,k} \cap Z$ where Z is a basic Zariski set of spec(R).*

(We have identified the point in spec(R) determined by values (t') of (t) in k.)

Proof. Without loss of generality, we may assume that k is infinite. Say $S = R[x_1, \cdots, x_n]$. Put

$$y = c_1x_1 + \cdots + c_nx_n$$

where the c_i are elements of k such that if $K = K(R)$ then $K(x_1, \cdots, x_n) = K(y)$. Such elements can be found by the usual arguments used to prove the primitive element theorem of Galois theory. Then spec(S) and spec$(R[y])$ contain a common Zariski basic set, and from this, taking for f an irreducible equation of y over R, our assertion is obvious.

Our rings S over R may be viewed as defining algebraic systems of 0-cycles, whose generic members are relatively irreducible over the quotient field of R. The Hilbert sets then give points of the parameter scheme over which special members of the family remain relatively irreducible. It is an easy matter (taking generic projections, as with Chow forms) to extend the definition to algebraic systems of arbitrary dimension, i.e. to assume merely that we are dealing with a scheme S over spec(R) which is of finite type, without divisors of zero, and whose generic member is irreducible as a scheme. One would see that the set of $\mathfrak{p} \in$ spec(R) for which the localized fiber $S_{\mathfrak{p}}$ is without divisors of zero and is irreducible over $k(\mathfrak{p})$ contains a Hilbert subset of spec(R).

One should not confuse this relative irreducibility with absolute irreducibility. For the convenience of the reader, let us recall here the standard result concerning the preservation of absolute irreducibility.

PROPOSITION 7. *Let R be a ring, K its quotient field, and $f(X_1, \cdots, X_s) \neq 0$ a polynomial in $R[X]$ which is irreducible in $\bar{K}[X]$. Then there exists $\xi \in R$, $\xi \neq 0$ such that for any $\mathfrak{p} \notin$ spec(R_ξ) the* *reduced polynomial* $f_{\mathfrak{p}}(X)$ *is irreducible in* $k(\mathfrak{p})[X]$.

Proof. Let f be of degree d. Consider two integers $m, n \geqq 1$ such that $m + n = d$. Let M_α (resp. M_β) range over all monomials of degree $\leqq m$ (resp. n) and put

$$g_u = \sum u_\alpha M_\alpha(X)$$
$$h_v = \sum v_\beta M_\beta(X)$$

where the u_α, v_β are a set of algebraically independent quantities over K. Then

$$g_u h_v = \sum \varphi_\lambda(u, v) M_\lambda(X)$$

where the $\varphi(u, v)$ are polynomials with integer coefficients and are "universal." We can write

$$f(X) = \sum c_\lambda M_\lambda(X) \qquad c_\lambda \in R\,.$$

By assumption, and Hilbert's nullstellensatz, the ideal generated by the $\varphi_\lambda(u, v) - c_\lambda$ in $K[u, v]$ is the unit ideal, and 1 has a finite expression in terms of these generators. Multiplying the cofficients by some element of R, we get

$$\xi = \sum \psi_\lambda[\varphi_\lambda(u, v) - c_\lambda]$$

with $\xi \in R$, $\xi \neq 0$, and $\psi_\lambda \in R[u, v]$.

For each m, n as above, we get such a relation, and we should actually index ξ by m, n, i.e. write $\xi_{m,n}$. Taking the product of these $\xi_{m,n}$ for each pair m, n together with the non-zero cofficients of f, we get an element ξ of R which will obviously satisfy our requirements.

§ 6. *Applications to commutative group varieties*

Let R be an integrally closed ring, and A a commutative group variety scheme over spec(R), or over R as we shall say. *We assume that it is simple over R* (this amounts to the classical terminology of non-degeneracy). The only properties of such schemes which we shall use are the following:

If $S_1 \subset S_2$ are two rings containing R, and if $A(S_1)$, $A(S_2)$ denote the points of A in S_1 and S_2, respectively, then the natural map

$$A(S_1) \to A(S_2)$$

is an injection. We identity $A(S_1)$ in $A(S_2)$.

For each $\mathfrak{p} \in \operatorname{spec}(R)$, the fiber $A_{\mathfrak{p}}$ over $\mathfrak{p}$ is a commutative group variety scheme over the residue class field $k(\mathfrak{p})$. If $x \in A(R)$ and $x(\mathfrak{p})$ denotes the corresponding point of $A_{\mathfrak{p}}$ in $k(\mathfrak{p})$, then the map

$$x \rightarrow x(\mathfrak{p})$$

is a homomorphism of $A(R)$ into $A_{\mathfrak{p}}(k(\mathfrak{p}))$.

If $S \supset R$, $\mathfrak{P} \in \operatorname{spec}(S)$, and m an integer >1, prime to the characteristic of $k(\mathfrak{P})$, then the map $x \rightarrow x(\mathfrak{P})$ defines an injection on the subgroup of points of $A(S)$ consisting of the points of period m.

If K is the quotient field of R, E a finite extension of K, and $a \in A(K)$, then there exists a subring S of E, finitely generated over R, such that $a \in A(S)$.

Given a point c' of period m in $A_{\mathfrak{p}}(k(\mathfrak{p}))$, there exists a point c of period m in some $A(S)$ for S finitely generated and algebraic over R such that $c' = c(\mathfrak{P})$ for some $\mathfrak{P} \in \operatorname{spec}(S)$ lying above $\mathfrak{p}$.

If $b \in A(E)$ for some extension E of K, then there exists a smallest field $K(b)$ of rationality for b containing K. If $b \in A(S)$, where $R \subset S \subset K(b)$ and $\mathfrak{P} \in \operatorname{spec}(S)$ lies above $\mathfrak{p}$, then $k(\mathfrak{P}) = k(\mathfrak{p})(b(\mathfrak{P}))$.

LEMMA 2. *Let R be integrally closed and A a commutative group variety scheme over R. Let $a \in A(R)$ and let m be an integer >0 prime to the characteristic of K. Suppose that $a \notin mA(K)$. Then there exists a ring S, finitely generated and algebraic over R, such that for all $\mathfrak{p} \in U_{S/R}$, the point $a(\mathfrak{p})$ does not lie in $mA_{\mathfrak{p}}(k(\mathfrak{p}))$.*

Porof. Since $a \notin mA(K)$, the field $K(b)$, for any point $b \in A(\bar{K})$ such that $mb = a$, has degree >1 over K. We may assume that all these points lie in $A(S)$ where S is finitely generated and algebraic over R. Replacing R by $R[1/\xi]$ for a suitable $\xi \neq 0$ in R, we may assume that S in finite over R. We also adjoint $1/m$ to R. We see then that for $\mathfrak{p} \in U_{S/R}$ and $\mathfrak{P} \in \operatorname{spec}(S)$ lying above $\mathfrak{p}$, the extension

$$k(\mathfrak{p})(b(\mathfrak{P})) \qquad \text{of } k(\mathfrak{p})$$

has degree >1. By hypothesis, the map $b \rightarrow b(\mathfrak{P})$ induces a bijection of the m-th roots of a onto the m-th roots of $a(\mathfrak{p})$. Hence $a(p)$ cannot lie in $mA_{\mathfrak{p}}(k(\mathfrak{p}))$.

If Γ is an abelian group, we denote by Γ_m the subgroup of points of period m.

THEOREM 3. *Let R be an integrally closed ring, and A a com-*

mutative group variety scheme, simple over R. Let m be an integer >1, prime to the characteristic of the quotient field K of R. Let Γ be a finitely generated subgroup of $A(R)$. Assume:

(i) *The period of each torsion element of Γ divides m.*

(ii) *The canonical map $\Gamma'/m\Gamma' \to A(K)/mA(K)$ is an injection (denoting by Γ' the group $\Gamma + A(K)_m$). Then there exists a Hilbert subset U of spec(R) such that for $\mathfrak{p} \in U$, the map*

$$x \to x(\mathfrak{p})$$

induces an injection on Γ.

Proof. Extending R by a finite number of elements of type $1/\xi$ with $\xi \neq 0$ in R, we may assume that $1/m \in R$ and that all points in $A(K)$ of period m lie in $A(R)$. We may then assume, without loss of generality, that these points are contained in Γ' (considering Γ' instead of Γ).

We can write Γ as a direct sum of a free abelian group B and a finite torsion group C. Let $x_1, \cdots, x_n$ be a basis for B over $\mathbf{Z}$. Let $\{a_j\}$ be a set of representatives in B of the factor group B/mB, whose coordinates in $\mathbf{Z}$ are the integers α such that $0 \leqq \alpha \leqq m-1$ or $0 \leqq -\alpha \leqq m-1$. We consider the finite set of points $a_j + c$, $c \in C$. By assumption, we can determine for each such point a Hilbert set as in the lemma. Further, we can find a ring S finitely generated, algebraic over R, such that every point $z \in A(\bar{K})$ with $mz = 0$ is in $A(S)$. This condition implies in particular that if $\mathfrak{p} \in U_{S/R}$ and $K(z) \neq K$ then $z(\mathfrak{p})$ is not rational over $k(\mathfrak{p})$. We let U be the intersection of the finite number of Hilbert sets just mentioned.

Let $\mathfrak{p} \in U$ and let $x \in \Gamma$ be such that $x(\mathfrak{p}) = 0$. To show that $x = 0$. we may assume that $x \in B$ and that $x(\mathfrak{p})$ has period m and derive a contradiction. By construction, $x(\mathfrak{p})$ cannot be equal to $z(\mathfrak{p})$, where z is a point of period m in $A(\bar{K})$ and not in $A(K)$. Hence it must be of type $c(\mathfrak{p})$ with $c \in C$. Consequently,

$$x(\mathfrak{p}) - c(\mathfrak{p}) = 0 .$$

For some a_j and some $y \in B$, we have

$$x = a_j + my$$

and hence

$$0 = a_j(\mathfrak{p}) - c(\mathfrak{p}) + my(\mathfrak{p}) .$$

By construction, $a_j - c = 0$, and hence $y(\mathfrak{p})$ has period m.

If we write $x = \lambda_1 x_1 + \cdots + \lambda_n x_n$ with $\lambda_i \in \mathbf{Z}$, then we can express y in terms of our basis $x_1, \cdots, x_n$ in the form

$$y = \mu_1 x_1 + \cdots + \mu_n x_n$$

where $m\mu_i = \lambda_i - \alpha_i$, the α_i being the coordinates of a_j with respect to our basis. If $x \neq 0$, then repeating the procedure a finite number of times, we get $a_j(\mathfrak{p}) + c(\mathfrak{p}) = 0$ for some a_j and some $c \in C$, which gives us our contradiction.

Corollary. *Let A_0 be an abelian variety defined over a field K, of finite type over $\mathbf{Q}$. Then there exists an integrally closed ring R of finite type over $\mathbf{Z}$ having K as quotient field, and an abelian variety scheme A over R such that A_0 is the extension of A to the ground field K, and $A(R) = A(K)$. There exist infinitely many prime ideals $\mathfrak{p}$ of R such that the residue class field $k(\mathfrak{p})$ is finite over $\mathbf{Q}$, and that the map*

$$x \to x(\mathfrak{p})$$

of $A(R)$ into $A_{\mathfrak{p}}(k(\mathfrak{p}))$ is an injection on $A(R)$. The above assertions remain true if A_0 is a complete non-singular curve of genus $\geqq 1$, and A is a scheme of such curves.

Proof. The first assertion is standard, and we can also make R integral over a ring $I[t_1, \cdots, t_r]$ where I is a ring generated over $\mathbf{Z}$ by $1/p_1, \cdots, 1/p_m$ for a finite number of primes p_i, and $t_1, \cdots, t_r$ are algebraically independent over $\mathbf{Q}$ (using the Noether normalization theorem or Proposition 2 of Chapter I, § 1). We can now use Propositions 5 and 6, the latter to apply Hilbert's irreducibility theorem. The generalized Mordell-Weil theorem (Theorem 1 of Chapter V) gives us the finite generation of $A(K)$, and consequently we can also apply Theorem 3, in the case of abelian varieties. In the case of curves of genus $\geqq 1$, we select R large enough so that a canonical map of the curve into its Jacobian is defined over R, and that the family of such maps is non-degenerate. Then the points of the curve may be viewed as a subset of the points of its Jacobian, and our assertion is reduced to the case of abelian varieties.

By means of this corollary, one can reduce certain qualitative assertions concerning points of curves over finitely generated rings and fields over $\mathbf{Q}$ to assertions about points on curves in number fields. For instance, the finiteness of integral points proved in

Chapter VII, for fields of finite type, is thus seen to be implied by the finiteness of integral points over finitely generated subrings of number fields over **Z**, *provided* one has Theorem 1 of Chapter V. In a like manner, the Mordell conjecture could be so reduced, if one could prove it in number fields.

Historical Note:

The proof of Hilbert's theorem over the rationals is due to Dörge [9], from which all of § 2 has been more or less copied. The reduction steps occur in Hilbert's original paper [12].

Siegel [37[had already observed the effect of his theorem on Hilbert's. The idea of counting rational points on curves of genus $\geqq 1$ and finding that their number is much smaller than the number of points on curves of genus 0 is due to Néron [24].

The proof of the function field case is due to Franz [10].

The theorem of § 6 concerning the possibility of specializing a finitely generated subgroup of an abelian variety injectively is due to Néron [24]. By this method, he gives examples of elliptic curves having non-trivial groups of rational points over certain number fields. It is unknown whether one can construct elliptic curves over the rationals such that the group of rational points has arbitrarily high rank. There is a general feeling that this connot be done. One could try to investigate the geometric case of a rational pencil of elliptic curves on a surface, as Igusa does [14], and to raise the same question for surfaces.

It would also be interesting to investigates to what extent the specialization theorem remains valid for non-commutative groups. For instance, given a finitely presented group of matrices in characteristic 0, does it always have an injective specialization into a group of matrices all of whose coefficients are algebraic numbers?

BIBLIOGRAPHY

[1] E. ARTIN, *Algebraic numbers and algebraic functions,* Mimeographed notes, Princeton and New York Universities, 1950–51.

[2] ——— and G. WHAPLES, *Axiomatic characterization of fields by the product formula,* Bull. Am. Math. Soc. 51, No. 7 (1945) pp. 469–492.

[3] BOURBAKI, Commutative algebra, Hermann, Paris, 1961.

[4] J. W. CASSELS, *Arithmetic on curves of genus* 1, J. Reine Angew. Math. 202 (1959) pp. 52–99.

[5] C. CHABAUTY, *Sur les équations diophantionnes liées aux unités d'un corps de nombres algébriques fini,* these, Annali di Math. 17 (1938) pp. 127–168.

[6] ———, *Sur les points rationnels des variétés algébriques dont l'irrégularité est supérieure à la dimension,* Compt. Rend. 212 (1941) pp. 1022–1024.

[7] H. DAVENPORT, *On a principle of Lipschitz,* J. London Math. Soc. 26 (1951) pp. 179–183.

[8] H. DAVENPORT and K. F. ROTH, *Rational approximations to algebraic numbers,* Mathematika 2 (1955) pp. 160–167.

[9] K. DÖRGE, *Einfacher Beweis des Hilbertschen Irreduzibilitätssätzes,* Math. Ann. 96 (1927) pp. 176–182.

[10] W. FRANZ, *Untersuchungen zum Hilbertschen Irreduzibilitätssatz,* Math. 33 (1931) pp. 275–293.

[11] A. GROTHENDIECK, and J. DIEUDONNÉ, Éléments de Géometrie algébrique, Publications Mathématiques de l'Institut des Hautes Études Scientifiques, Paris, 1960.

[12] D. HILBERT, *Uber die Irreduzibilität ganzer rationaler Funktionen mit ganzzahligen Koeffizienten,* J. Reine Angew. Math. 110 (1892) pp. 104–129.

[13] E. INABA, *Uber den Hilbertschen Irreduzibilitätssatz,* Jap. J. Math. 19 (1944) pp. 1–25.

[14] J. IGUSA, *Betti and Picard numbers of algebraic surfaces,* Proc. Nat. Acad. Sci. U.S.A. 45, No. 5 (1960) pp. 724–726.

[15] S. LANG, *Unramified class field theory over function fields in several variables,* Ann. of Math. 64, No. 2 (1956) pp. 285–325.

[16] ———, Introduction to algebraic geometry, Interscience, New York, 1959.

[17] ———, Abelian varieties, Interscience, New York, 1959.

[18] ———, *Integral points on curves,* Publications mathématiques de l'Institut des Hautes Études Scientifiques, Paris, 1960.

[19] ——— and A. NÉRON, *Rational points of abelian varieties over function fields,* Am. J. Math. 81, No. 1 (1959) pp. 95–118.

[20] ——— and J. TATE, *Principal homogeneous spaces over abelian varieties,* Am. J. Math. 80 (1958) pp. 659–684.

[21] K. MAHLER, *Uber die rationalen Punkte auf Kurven vom Geschlecht Eins,* J. Reine Angew. Math. 170 (1934) pp. 168–178.

[22] ———, *On the fractional parts of the powers of a rational number,* Mathematika, 4 (1957) pp. 122–124.

[23] L. J. MORDELL, *On the rational solutions of the indeterminate equation of the third and fourth degrees,* Proc. Cambridge Philos. Soc. 21 (1922).

[24] A. NÉRON, *Problèmes arithmétiques et géométriques rattachés à la notion de rang d'une courbe algébrique dans un corps,* Bull. Soc. Math. France 80 (1952) pp. 101-166.

[25] ———, *Arithmétique et classes de diviseurs sur les variétés algébriques,* Proc. International Symposium on Algebraic number theory, Tokyo-Nikko, Tokyo, 1955, pp. 139–154.

[26] D. G. NORTHCOTT, (a) *An inequality in the theory of arithmetic on algebraic varieties,* Proc. Cambridge Philos. Soc. 45 (1949) pp. 502–509. (b) *A further inequality in the theory of arithmetic on algebraic varieties,* Proc. Cambridge Philos. Soc. 45 (1949) pp. 510-518.

[27] D. RIDOUT, *The p-adic generalization of the Thue-Siegel-Roth Theorem,* Mathematika 5 (1958) pp. 40–48.

[28] ———, *Rational approximations to algebraic numbers,* Mathematika 4 (1957) pp. 125–131.

[29] P. ROQUETTE, *Uber das Hassesche Klassenkörper Zerlegungsgesetz und seine Verallgemeinerung fur beliebige abelsche Funktionenkörper,* J. Reine Angew. Math. 197 (1957) pp. 49–67.

[30] ———, *Einheiten und Divisorenklassen in endlich erzeugbar Körpern,* J. Deutsch. Math. Verein. 60 (1958) pp. 1–27.

[31] K. F. ROTH, *Rational approximations to algebraic numbers,* Mathematika 2 (1955) pp. 1–20.

[32] T. SCHNEIDER, Einführung in die Théorie der transcendental Zahlen, Springer Verlag, Berlin, 1957.

[33] J. P. SERRE, Groupes algébriques et théorie du corps de classes, Hermann, Paris, 1958.

[34] I. ŠAFAREVIČ, *The group of principal homogeneous algebraic manifolds,* Doklady Akad. Nauk SSSR 124 (1959) pp. 42–43.

[35] ———, *Exponents of Eliptic curves,* Doklady Akad. Nauk SSSR 114 (1957) pp. 714–716.

[36] G. SHIMURA, *Reduction of algebraic varieties with respect to a discrete valuation of the basic field,* Am. J. Math. 77 (1955) pp. 134–176.

[37] G. L. SIEGEL, *Über einige Anwendungen Diophantischer Approximationen,* Abh. Preuss. Akad. Wiss. Phys. Math. Kl. (1929) pp. 41–69.

[38] J. TATE, WC groups over p-adic fields, Seminaire Bourbaki, 1958.

[39] VAN DER WAERDEN, Moderne Algebra, Teubner, Berlin, 1930.

[40] A. WEIL, *Arithmetic on algebraic varieties,* Ann. of Math. 53, Vo. 3 (1951) pp. 412–444.

[41] A. WEIL, *L'arithmétique sur les courbes algébriquss,* Acta Math. 52 (1928) pp. 281–315.

[42] ———, Foundations of Algebraic Geometry, New York, second edition.

[43] ———, *Sur les critères d'équivalences en géometrie algébrique,* Math. Ann. 128 (1954) pp. 95–127.

TABLE OF NOTATION

I include here the notation used most frequently, together with a short informal description of each term.

R^*, K^*	Invertible elements in a ring R (resp. non-zero elements of a field K).
$\bar{K}$	Algebraic closure of a field K.
$K(P)$	Field obtained by adjoining to K a set of affine coordinates for a point P (equal to $K(x_0/x_i, \cdots, x_n/x_i)$ if $(x_0, \cdots, x_n)$ are projective coordinates for P, $x_i \neq 0$).
$\mathbf{N}\mathfrak{p}$	Number of elements in $\mathfrak{o}/\mathfrak{p}$.
$\mathbf{N}\mathfrak{a}$	Extended by linearity from $\mathbf{N}\mathfrak{p}$.
N_K^E	Norm from E to K.
V_K, A_K	Set of rational points of a variety over K.
M_K	Proper set of absolute values.
$\lvert\mathfrak{d}\rvert_v$	$\mathfrak{d}(v)$ if $\mathfrak{d}$ is an M_K-divisor.
$\lVert\mathfrak{d}\rVert_v$	$\lvert\mathfrak{d}\rvert_v^{N_v}$.
$\lVert\mathfrak{d}\rVert_K$	The product $\Pi_{v\in M_K}\lVert\mathfrak{d}\rVert_v$.
h_φ, H_φ	Height determined by the map φ into projective space.
H, H_K	Relative height to a field K.
h_X, H_X	Height determined by a mapping derived from the linear system $\mathscr{L}(X)$.
$\sim$	Linear equivalence, or equivalence of functions (each is less than a constant multiple of the other).
$\approx$	Algebraic equivalence, or quasi-equivalence of functions (same as above, but with $1+\varepsilon$ and $1-\varepsilon$ in the exponent).

INDEX

Absolute height, p. 44
Absolute value, p. 1
Algebraic (ring over another), p. 155
Attach a distribution, p. 130
Attach rationally, p. 132

Basic Hilbert set, p. 155
Basic Zariski set, p. 155
Belong (in Kummer theory), p. 78

Canonical set of absolute values, p. 21
Completion, p. 6
Complex absolute value, p. 21
Complex prime, p. 21
Constants, p. 92
Constants for derivations, p. 106

Defined at a point (distribution), p. 129
Degree (of a monomial), p. 106
Degree of a prime, p. 23
Dependent absolute values, p. 1
Differential operator, p. 108
Differential ring, p. 106
Discrete, p. 4
Distribution, p. 128

Equivalent distributions, p. 130
Equivalent functions, p. 61
Exponent (of Kummer extensions), p. 75
Extend (prime), p. 3

Finite divisors, p. 26
Finite primes, p. 21
Function field, p. 23

Height of a point, p. 43
Height of a polynomial, p. 48
Hilbert set (abstract), p. 155
Hilbert subset of S_k^r, p. 141
Hilbertian, p. 141

Independent, p. 2
Integers, p. 91
Integral M_K-ideal, p. 27

Kronecker substitution, p. 150
Kummer extensions, p. 78

Lift back, p. 10
Local degree, p. 14
Local norm, p. 16
Logarithmic height, p. 50

M_K-divisor, p. 25
M_K-ideal, p. 26
M_K-ideal class, p. 26
M_K-unit, p. 27

Non-archimedean absolute value, p. 1
Non-archimedean prime, p. 3
Non-degenerate function, p. 55
Non-degenerate reduction, p. 80
Number field, p. 2

Order of a field element, p. 4
Order of an M_K-divisor, p. 26
Order of an M_K-ideal, p. 27

Prime, p. 3
Prime at infinity, p. 21
Prime divisor, p. 22
Prime divisor lying above $\mathfrak{p}$, p. 24
Principal M_K-ideal, p. 26
Principal M_K-divisor, p. 26

Product formula, p. 20
Product formula with multiplicities, p. 20
Proper absolute value, p. 19
Proper set of absolute values, p. 19

Quasi-degree, p. 55
Quasi-equivalent, p. 61

Ramification index, p. 3
Rational (distribution over a field), p. 129
Real absolute value, p. 21
Real prime, p. 21
Relative height of a point, p. 43
Relative height of a polynomial, p. 48
Residue class degree, p. 3
Root (of a point), p. 76

Size of a divisor, p. 30
S-integers, p. 91
S-units, p. 27

Torus, p. 134
Trivial absolute value, p. 1

Unramified set of absolute values, p. 78
Unramified valuation, p. 10

Valuation, p. 1
Variety scheme, p. 24

Well-behaved absolute value, p. 15

3 9000 022 490 531

T2-EQL-630